A Strong Order 1/2 Method for Multidimensional SDEs with Discontinuous Drift

Gunther Leobacher and Michaela Szölgyenyi

December 2015

Abstract

In this paper we consider multidimensional stochastic differential equations (SDEs) with discontinuous drift and possibly degenerate diffusion coefficient. We prove an existence and uniqueness result for this class of SDEs and we present a numerical method that converges with strong order 1/2. Our result is the first one that shows strong convergence for such a general class of SDEs.

The proof is based on a transformation technique that removes the discontinuity from the drift such that the coefficients of the transformed SDE are Lipschitz continuous. Thus the Euler-Maruyama method can be applied to this transformed SDE. The approximation can be transformed back, giving an approximation to the solution of the original SDE.

As an illustration, we apply our result to an SDE the drift of which has a discontinuity along the unit circle.

Keywords: stochastic differential equations, discontinuous drift, degenerate diffusion, numerical methods for stochastic differential equations, strong convergence
Mathematics Subject Classification (2010): 60H10, 65C30, 65C20 (Primary), 65L20 (Secondary)

1 Introduction

We consider a d-dimensional time-homogeneous stochastic differential equation (SDE)

$$dX = \mu(X)\,dt + \sigma(X)dW\,, \qquad X_0 = x\,, \tag{1}$$

where $\mu : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^{d\times d}$ are measurable functions and $W = (W_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

If both μ and σ are Lipschitz, then existence and uniqueness is guaranteed by Picard iteration. Furthermore, (1) can be solved numerically with, e.g., the Euler-Maruyama method, which then converges with strong order 1/2, see [9, Theorem 10.2.2].

However, in applications one is frequently confronted with SDEs where μ is non-Lipschitz, e.g., in stochastic control theory. The question of existence and uniqueness of solutions to SDEs with non-Lipschitz drift has been studied by various authors.

For the case where μ is only bounded and measurable and σ is bounded, Lipschitz, and satisfies a certain uniform ellipticity condition, Zvonkin [23] and Veretennikov [20, 21, 22] prove existence and uniqueness of a solution by locally removing the drift coefficient in a way such that locally the Lipschitz condition of the diffusion coefficient is preserved.

But uniform ellipticity is a strong assumption which is often violated in applications.

In Leobacher et al. [14] an existence and uniqueness result for (1) is presented for the case where the drift is potentially discontinuous at a hyperplane, or a special hypersurface, but well behaved everywhere else and where the diffusion coefficient is potentially degenerate. In that paper, not the whole drift is removed, but only the discontinuity is removed locally from the drift.

Due to the weaker requirements on the diffusion coefficient the restriction to homogeneous SDEs does not pose any loss of generality. In Shardin and Szölgyenyi [17] the authors extend the result from [14] to the time-inhomogenous case.

In Leobacher and Szölgyenyi [12] an existence and uniqueness result, as well as a numerical method are presented for the one-dimensional case with piecewise Lipschitz drift coefficient. There the coefficients are globally transformed into Lipschitz ones. Both computation of the transformed coefficients and inversion can be done efficiently. This leads to a numerical method for one-dimensional SDEs through application of the Euler-Maruyama scheme on the transformed equation and transforming the approximation back. We give a simplified version of this result in Section 2.

However, extending the result from [12] to the d-dimensional case is not straightforward. One problem is that there is no immediate generalization of the concept of a piecewise Lipschitz function with several variables that suites our needs. The second problem is that it is more difficult to obtain a transform that is a Lipschitz diffeomorphism $\mathbb{R}^d \longrightarrow \mathbb{R}^d$. We use Hadamard's global inverse function theorem to prove that our transform is of this kind. Moreover, we need to show that the transform and its inverse are sufficiently well-behaved for Itô's formula to hold.

The coefficients of the SDE obtained by transforming the original one are shown to be Lipschitz, such that we can apply the Euler-Maruyama method to the transformed SDE and obtain the (best possible) convergence order $1/2$. An approximation to the original SDE is then obtained by applying the inverse transform to the approximation of the transformed solution. One might ask whether the results of Zvonkin and Veretennikov give rise to a similar method. However, to apply their method one would have to solve a system of parabolic partial differential equations in each step. Further, for using this solution in a numerical method like ours, one would also have to find its inverse function. Therefore such a method, if it exists at all, would be rather costly from the computational perspective.

In the present paper we present a transform for the multidimensional case which allows to prove an existence and uniqueness result for d-dimensional SDEs with discontinuous drift and degenerate diffusion coefficient under conditions significantly weaker than those in the literature. The essential geometric condition in our setup is that the diffusion must have a component orthogonal to the set of discontinuities of the drift.

Furthermore, we present a numerical method for such SDEs based on the ideas outlined above. Up to the authors' knowledge there is no other numerical method that can deal with such a general class of SDEs and gives strong convergence.

We are now going to review the literature on numerical methods for SDEs with non-globally Lipschitz drift coefficient. In Berkaoui [1] strong convergence of the Euler-Maruyama scheme is proven under the assumption that the drift is of class C^1. For an SDE with continuously differentiable but non-globally Lipschitz drift Hutzenthaler et al. [7] introduce a new explicit numerical scheme – the tamed Euler scheme – and prove its strong convergence. Sabanis [16] proves strong convergence of the tamed Euler scheme for SDEs with one-sided Lipschitz drift. For the Euler-Maruyama scheme Gyöngy [5] proves almost sure convergence for the case that the drift satisfies a monotonicity condition. A different approach is introduced by Halidias and Kloeden [6], who show that the Euler-Maruyama scheme converges strongly for SDEs with a discontinuous monotone drift coefficient, especially mentioning the case in which the drift is a Heaviside function. Kohatsu-Higa et al. [10] show weak convergence of a method where they first regularize the drift and then apply the Euler-Maruyama scheme. They allow the drift to be discontinuous. Étoré and Martinez [2, 3] introduce an exact simulation algorithm for SDEs that have a bounded drift coefficient being discontinuous in one point, but differentiable everywhere else.

This paper is organized as follows. In Section 2 we give the one-dimensional result and algorithm in a form that can be generalized to multiple dimensions, which subsequently is done in Section 3.

In Section 4 we give a numerical example for the purpose of illustration, where the drift coefficient has discontinuities along the unit circle in $\mathbb{R}^2$.

2 The one-dimensional problem

Here we consider the one-dimensional version of SDE (1).

First we assume that $\mu : \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz only with the exception of a single point where μ is allowed to have a jump. W.l.o.g. this point is 0. Thus we have

Assumption 2.1. For $\mu\colon \mathbb{R} \longrightarrow \mathbb{R}$, let $\mu|_{\mathbb{R}^+}, \mu|_{\mathbb{R}^-}$ be Lipschitz.

Assumption 2.2. Let $\sigma : \mathbb{R} \longrightarrow \mathbb{R}$ be Lipschitz with $\sigma(0) \geq c_0 > 0$.

We are going to construct a transform $G : \mathbb{R} \longrightarrow \mathbb{R}$ such that the process formally defined by $Z = G(X)$ satisfies an SDE with Lipschitz coefficients and therefore has a solution by the classical Itô theorem on existence and uniqueness of solutions.

For this define the following bump function on $\mathbb{R}$, which we need to localize the impact of the transform G:

$$\phi(u) = \begin{cases} (1+u)^3(1-u)^3 & \text{if } |u| \leq 1\,, \\ 0 & \text{else}\,. \end{cases}$$

The function ϕ has the following properties:

1. ϕ defines a C^2 function on all of $\mathbb{R}$;
2. $\phi(0) = 1$, $\phi'(0) = 0$, $\phi''(0) = -6$;
3. $\phi(\pm 1) = \phi'(\pm 1) = \phi''(\pm 1) = 0$.

We define the transform $G : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$G(x) = x + \alpha\phi\left(\frac{x}{c}\right) x|x|\,, \quad x \in \mathbb{R}\,, \tag{2}$$

where α and $c > 0$ are some constants.

Lemma 2.1. *Let* $c < \frac{1}{6|\alpha|}$.

Then $G'(x) > 0$ *for all* $x \in \mathbb{R}$. *Furthermore,* $G'(x) = 1$ *for all* $|x| > c$. *Therefore* G *has a global inverse* G^{-1}.

Proof. Differentiating G yields

$$G'(x) = 1 - \frac{6\alpha x^2|x|}{c^2}\left(1+\frac{x}{c}\right)^2\left(1-\frac{x}{c}\right)^2 + 2\alpha|x|\left(1+\frac{x}{c}\right)^3\left(1-\frac{x}{c}\right)^3 .$$

For positive α this is positive, if $c < \frac{1}{6|\alpha|}$. For negative α it is positive, if $c < \frac{1}{2|\alpha|}$. Altogether a sufficient condition for G' to be positive is $c < \frac{1}{6|\alpha|}$. □

W.l.o.g. we always choose $c < \frac{1}{6|\alpha|}$, such that G has a global inverse.

Remark 2.2. In [12] the function G is constructed differently. There, G is piecewise cubic, such that G^{-1} is piecewise radical and hence admits exact inversion, which is advantageous for the numerical treatment.

In fact, G can be made piecewise cubic by still using equation (2), but with a different choice for ϕ. Actually, any function ϕ with support contained in $[-1, 1]$ satisfying properties 1., 2., 3. from above will give rise to a transform G sufficient for our purpose.

The form chosen here is simple in the one-dimensional case and has a direct multidimensional analog.

Formally define $Z = G(X)$. Abbreviating $\bar{\phi}(x) := \phi(\frac{x}{c})x|x|$, we have

$$\begin{aligned} dZ &= dX + \alpha\bar{\phi}'(X)dX + \frac{1}{2}\alpha\bar{\phi}''(X)d[X] && (3)\\ &= \left(\mu(X) + \alpha\bar{\phi}'(X)\mu(X) + \frac{1}{2}\alpha\bar{\phi}''(X)\sigma(X)^2\right)dt + \left(\sigma(X) + \alpha\bar{\phi}'(X)\sigma(X)\right)dW && (4)\\ &= \tilde{\mu}(Z)dt + \tilde{\sigma}(Z)dW\,, && (5) \end{aligned}$$

where

$$\begin{aligned} \tilde{\mu}(z) &= \mu(G^{-1}(z)) + \alpha\bar{\phi}'(G^{-1}(z))\mu(G^{-1}(z)) + \frac{1}{2}\alpha\bar{\phi}''(G^{-1}(z))\sigma(G^{-1}(z))^2\,,\\ \tilde{\sigma}(z) &= \sigma(G^{-1}(z)) + \alpha\bar{\phi}'(G^{-1}(z))\sigma(G^{-1}(z))\,. \end{aligned}$$

We now show that, for an appropriate choice of α, the transformed drift $\tilde{\mu}$ is Lipschitz. For this we need the following definition and lemma from [12].

Definition 2.3. Let $I \subseteq \mathbb{R}$ be an interval. We say a function $f : I \longrightarrow \mathbb{R}$ is piecewise Lipschitz, if there are finitely many points $a_1 < \ldots < a_m \in I$ such that f is Lipschitz on each of the intervals $(-\infty, a_1) \cap I$, $(a_m, \infty) \cap I$ and (a_k, a_{k+1}), $k = 1, \ldots, m$.

Lemma 2.4. *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be piecewise Lipschitz and continuous.*
Then f is Lipschitz on $\mathbb{R}$.

Proof. The elementary proof can be found in [12]. □

From Lemma 2.4 and $\lim_{h\to 0}\bar{\phi}'(h) = 0$ we see that the mapping $z \mapsto \bar{\phi}'(G^{-1}(z))\mu(G^{-1}(z))$ is Lipschitz. In order to make the mapping $z \mapsto \mu(G^{-1}(z)) + \frac{1}{2}\alpha\bar{\phi}''(G^{-1}(z))\sigma(G^{-1}(z))^2$ continuous, we need to choose α so that

$$\mu(G^{-1}(0+)) + \frac{1}{2}\alpha\bar{\phi}''(G^{-1}(0+))\sigma(G^{-1}(0+))^2 = \mu(G^{-1}(0-)) + \frac{1}{2}\alpha\bar{\phi}''(G^{-1}(0-))\sigma(G^{-1}(0-))^2\,,$$

i.e.,

$$\mu(0+) + \frac{1}{2}\alpha\bar{\phi}''(0+)\sigma(0)^2 = \mu(0-) + \frac{1}{2}\alpha\bar{\phi}''(0-)\sigma(0)^2\,.$$

Thus we get, for the choice

$$\alpha = -2\frac{\mu(0+) - \mu(0-)}{(\bar{\phi}''(0+) - \bar{\phi}''(0-))\sigma(0)^2} = \frac{\mu(0-) - \mu(0+)}{2\sigma(0)^2}$$

that $\tilde{\mu}$ is continuous. Note that at this point we need non-degeneracy of σ in 0.

Since $\tilde{\mu}$ is continuous with the appropriate choice of α, it is Lipschitz as well by Lemma 2.4.

One may worry about the quadratic occurrence of σ in the expression for $\tilde{\mu}$. Note, however, that $\bar{\phi}''$ vanishes outside $[-c, c]$.

To prove that $\tilde{\sigma}$ is Lipschitz as well, we need the following lemma:

Lemma 2.5. *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be Lipschitz. Then $f\phi'$ is Lipschitz.*

Proof. Let L_f be a Lipschitz constant for f. Note that 6 is a Lipschitz constant for ϕ'. If $|x|, |y| \leq 1$, then

$$\begin{aligned} |f(x)\phi'(x) - f(y)\phi'(y)| &\leq |f(x)\phi'(x) - f(y)\phi'(x)| + |f(y)\phi'(x) - f(y)\phi'(y)| \\ &\leq L_f|x-y| \max_{z\in[-1,1]} |\phi'(z)| + 6|x-y| \max_{z\in[-1,1]} |f(z)| \\ &\leq K|x-y|\,, \end{aligned}$$

where $K = L_f|x-y| \max_{z\in[-1,1]} |\phi'(z)| + 6|x-y| \max_{z\in[-1,1]} |f(z)|$. For $-1 \leq x \leq 1 < y$ we have

$$\begin{aligned} |f(x)\phi'(x) - f(y)\phi'(y)| = |f(x)\phi'(x)| &= |f(x)\phi'(x) - f(1)\phi'(1)| \\ &\leq K|x-1| \leq K|x-y|\,. \end{aligned}$$

For $|x|, |y| > 1$ we have $|f(x)\phi'(x) - f(y)\phi'(y)| = 0 \leq K|x-y|$. ☐

Thus, $\tilde{\sigma}$ is Lipschitz by Lemma 2.5 and the fact that the composition of Lipschitz functions is Lipschitz.

Altogether we have that the SDE (3) for Z has Lipschitz coefficients $\tilde{\mu}$ and $\tilde{\sigma}$.

We are ready to prove existence and uniqueness of a solution to the one-dimensional SDE (1).

Theorem 2.6 (cf. [19, Theorem 2.2])**.** *Let Assumptions 2.1, and 2.2 be satisfied, i.e., μ is piecewise Lipschitz with at most one jump at 0, σ is Lipschitz and does not vanish in the discontinuity of μ, and G is globally invertible.*

Then the one-dimensional SDE (1) *has a unique global strong solution.*

Proof. Since the SDE (3) for Z has Lipschitz coefficients, it follows that (3) with initial condition $Z_0 = G(x)$ has a unique global strong solution for any $x \in \mathbb{R}$. Furthermore, G has a global inverse G^{-1}, which inherits the smoothness from G. Although $G^{-1} \notin C^2$, Itô's formula holds for G^{-1}, see [8, 5. Problem 7.3]. Applying Itô's formula to G^{-1}, we obtain that $G^{-1}(Z)$ satisfies

$$dX = \mu(X)dt + \sigma(X)dW\,, \quad X_0 = x\,.$$

Setting $X = G^{-1}(Z)$ yields the desired result. ☐

For approximating the solution to the one-dimensional SDE (1) we propose the following numerical method. Let $Z_T^{(\delta)}$ be the Euler-Maruyama approximation of the solution to SDE (3) with step size smaller than $\delta > 0$.

Algorithm 2.7. Go through the following steps:

1. Set $Z_0^{(\delta)} = G(x)$.
2. Apply the Euler-Maruyama method to (3) to obtain $Z_T^{(\delta)}$.
3. Set $\bar{X} = G^{-1}\left(Z_T^{(\delta)}\right)$.

Theorem 2.8 (cf.[19, Theorem 3.1])**.** *Let Assumptions 2.1, and 2.2 be satisfied.*

Then Algorithm 2.7 converges with strong order $1/2$ to the solution X of the one-dimensional SDE (1).

Proof. We estimate the L^2-error of the approximation. For every $T > 0$ there is a constant C, such that

$$\mathbb{E}\left(\left(X_T - \bar{X}_T\right)^2\right) = \mathbb{E}\left(\left(G^{-1}\left(Z_T\right) - G^{-1}\left(Z_T^{(\delta)}\right)\right)^2\right) \le L_{G^{-1}}^2 \mathbb{E}\left(\left(Z_T - Z_T^{(\delta)}\right)^2\right) = L_{G^{-1}}^2 C\delta$$

for every sufficiently small step size δ, where $L_{G^{-1}}$ is the Lipschitz constant of G^{-1} and where we applied [9, Theorem 10.2.2] for the L^2-convergence of the Euler-Maruyama scheme for SDEs with Lipschitz coefficients. □

3 The multidimensional problem

We now consider the multidimensional case. Like in the one-dimensional case, we will have to make assumptions on the drift so that it is Lipschitz apart from – relatively few – locations of discontinuity. That is, we need a concept similar to that of "piecewise Lipschitz" in the one-dimensional case. We will develop such a concept now.

In contrast to the one-dimensional case, we shall have to make additional assumptions on the behaviour of the drift close to its points of discontinuity, which shall all lie in a hypersurface Θ.

Regarding the diffusion coefficient we need to find a condition corresponding to Assumption 2.2.

Note that most of these assumptions are automatically satisfied, or can at least be weakened, if Θ is compact. We will treat the case of compact Θ in Section 3.6.

3.1 Piecewise Lipschitz functions

For a rectifiable curve $\gamma : [0, 1] \longrightarrow \mathbb{R}^d$, let $\ell(\gamma)$ denote its length,

$$\ell(\gamma) = \sup_{n, 0 \le t_1 < \ldots < t_n \le 1} \sum_{k=1}^{n} \|\gamma(t_k) - \gamma(t_{k-1})\| .$$

Definition 3.1. Let $A \subseteq \mathbb{R}^d$. The *intrinsic metric* d on A is given by

$$\rho(x, y) := \inf\{\ell(\gamma) : \gamma : [0, 1] \to A \text{ is a rectifiable continuous curve satisfying } \gamma(0) = x,\ \gamma(1) = y\} ,$$

where $\rho(x, y) := \infty$, if there is no rectifiable continuous curve from x to y.

Definition 3.2. Let $A \subseteq \mathbb{R}^d$. Let $f : A \longrightarrow \mathbb{R}^m$ be a function. We say that f is *intrinsic Lipschitz*, if it is Lipschitz w.r.t. the intrinsic metric on A, i.e. if there exists a constant L such that

$$\forall x, y \in A : \|f(x) - f(y)\| \le L\rho(x, y) .$$

Remark 3.3. Note that for a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ we have that f is piecewise Lipschitz, iff f is intrinsic Lipschitz on $\mathbb{R}\backslash A$, where A is a finite subset of $\mathbb{R}$.

This motivates the following definition:

Definition 3.4. We say that a function $f : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ is *piecewise Lipschitz*, if there exists a hypersurface[1] Θ with finitely many components and with the property, that the restriction $f|_{\mathbb{R}^d\backslash\Theta}$ is intrinsic Lipschitz. We call Θ the *exceptional set* for f.

The definition is more general than the more obvious requirement that $\mathbb{R}^d$ can be partitioned into finitely many patches in a way such that f is Lipschitz on all of the patches. This is illustrated by the following example.

[1]By a hypersurface we mean a $(d-1)$-dimensional submanifold of the $\mathbb{R}^d$.

Example 3.5. Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x) = \|x\| \arg(x)$. Then f is not Lipschitz, since $\lim_{h\to 0+} f(\cos(\pi - h), \sin(\pi - h)) = \pi$ and $\lim_{h\to 0+} f(\cos(\pi + h), \sin(\pi + h)) = -\pi$ for $x_1 < 0$.

It is readily checked, however, that f is intrinsic Lipschitz on $A = \mathbb{R}^2 \backslash \{x \in \mathbb{R}^2 : x_1 < 0, x_2 = 0\}$ and $\{x \in \mathbb{R}^2 : x_1 < 0, x_2 = 0\}$ is obviously a one-dimensional submanifold of $\mathbb{R}^2$.

Thus f is piecewise Lipschitz in the sense of Definition 3.4.

The following lemma is a multidimensional generalization of Lemma 2.4.

Lemma 3.6. *Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a function. If*

1. *f is continuous in every point $x \in \mathbb{R}^d$;*
2. *f is piecewise Lipschitz with exceptional set Θ;*
3. *for $x, y \in \mathbb{R}^d$ and $\eta > 0$ there exists a continuous rectifiable curve γ from x to y with $\ell(\gamma) < \|x - y\| + \eta$ such that $\#(\gamma \cap \Theta) < \infty$.*

Then f is Lipschitz on $\mathbb{R}^d$ w.r.t. the Euclidean metric, and with the same Lipschitz constant.

Proof. Let L be the intrinsic Lipschitz constant of f, i.e., $\|f(y) - f(x)\| \le L\rho(x, y)$ for all $x, y \in \mathbb{R}^d$, and let $x, y \in \mathbb{R}^d$. If $\rho(x, y) = \|x - y\|$, then clearly $\|f(x) - f(y)\| \le L\rho(x, y) = L\|x - y\|$.

If $\rho(x, y) > \|x - y\|$, then the line segment $s(x, y) := \{(1 - \lambda)x + \lambda y : \lambda \in [0, 1]\}$ has non-empty intersection with Θ.

Consider first the case where $s(x, y) \cap \Theta = \{z_1, \ldots, z_n\}$, i.e., we have finite intersection. There exist $\lambda_1, \ldots, \lambda_n$ such that $z_k = (1 - \lambda_k)x + \lambda_k y$. Define $g : [0, 1] \longrightarrow \mathbb{R}^m$ by $g(\lambda) := f((1 - \lambda)x + \lambda y)$.

Set $z_0 = x$, $z_{n+1} = y$, $\lambda_0 = 0$, $\lambda_{n+1} = 1$. W.l.o.g., $\lambda_0 < \ldots < \lambda_{n+1}$. Now

$$
\begin{aligned}
\|f(y) - f(x)\| = \|\sum_{k=1}^{n+1} f(z_k) - f(z_{k-1})\| \le \sum_{k=1}^{n+1} \|f(z_k) - f(z_{k-1})\| = \sum_{k=1}^{n+1} \|g(\lambda_k) - g(\lambda_{k-1})\| \\
&= \lim_{h\to 0+} \sum_{k=1}^{n+1} \|g(\lambda_k - h) - g(\lambda_{k-1} + h)\| \\
&\le \lim_{h\to 0+} \sum_{k=1}^{n+1} L\rho\Big(\big((1 - \lambda_k + h)x + (\lambda_k - h)y\big), \big((1 - \lambda_{k-1} - h)x + (\lambda_{k-1} + h)y\big)\Big) \\
&= \lim_{h\to 0+} \sum_{k=1}^{n+1} L\left\|\big((1 - \lambda_k + h)x + (\lambda_k - h)y\big) - \big((1 - \lambda_{k-1} - h)x + (\lambda_{k-1} + h)y\big)\right\| \\
&= \sum_{k=1}^{n+1} L\,\|z_k - z_{k-1}\| = L\|y - x\|\,,
\end{aligned}
$$

where we have used the continuity of f and g, and that the intrinsic metric coincides with the Euclidean metric for pairs of points for which the connecting line segment has empty intersection with Θ.

If $s(x, y) \cap \Theta$ contains infinitely many points, we can replace $s(x, y)$ by γ, which is only slightly longer than $s(x, y)$, but has only finitely many intersections with Θ. A slight modification of the argument above then gives that $\|f(y) - f(x)\| < L\|y - x\| + \varepsilon$ for any $\varepsilon > 0$, and thus the desired result. □

Conjecture 3.7. *Item 3 of the assumptions of Lemma 3.6 is not necessary to prove the assertion of the lemma.*

It is well-known that differentiable functions with bounded derivative are Lipschitz w.r.t. the euclidean metric. The same holds true for the intrinsic metric:

Lemma 3.8. *Let $A \subseteq \mathbb{R}^d$ be open and let $f : A \longrightarrow \mathbb{R}^m$ be differentiable with $\|f'\| \le K$.*

Then f is intrinsic Lipschitz with constant K.

Proof. Let $x, y \in A$ and let γ be a continuous curve of finite length with $\gamma(0) = x$ and $\gamma(1) = y$. (If no such curve exists we trivially have $\|f(y) - f(x)\| \le K\rho(x, y) = \infty$.) Let $0 = t_0 < \ldots < t_n = 1$. Without loss of generality the t_k can be chosen such that the line segment spanned by $\gamma(t_{k-1})$ and $\gamma(t_k)$ is in A für every k. Then

$$\begin{aligned}
\|f(y) - f(x)\| &\le \sum_{k=1}^{n} \|f(\gamma(t_k)) - f(\gamma(t_{k-1})\| \\
&\le \sum_{k=1}^{n} \sup_{t \in (t_{k-1}, t_k)} \|f'(\gamma(t)\| \, \|\gamma(t_k) - \gamma(t_{k-1})\| \\
&\le K \sum_{k=1}^{n} \|\gamma(t_k) - \gamma(t_{k-1})\| \le K\ell(\gamma)\,.
\end{aligned}$$

□

Furthermore, we prove that the composition of an intrinsic Lipschitz function with a Lipschitz function is intrinsic Lipschitz:

Lemma 3.9. *Let $A \subseteq \mathbb{R}^d$ be open. Let $g : \mathbb{R}^d \longrightarrow A$ be Lipschitz with constant L_g. Let $f : A \longrightarrow \mathbb{R}^m$ be intrinsic Lipschitz with constant L_f.*

Then $f \circ g$ is intrinsic Lipschitz with constant $L_f L_g$.

Proof. Let γ be a continuous curve of finite length with $\gamma(0) = x$ and $\gamma(1) = y$. (If no such curve exists we trivially have $\|f(y) - f(x)\| \le L_g \rho(x, y) = \infty$.) Let $0 = t_0 < \ldots < t_n = 1$. For every $\delta > 0$ there are $0 = \bar{t}_0 < \ldots < \bar{t}_{\bar{n}} = 1$ such that $\rho(g(x), g(y)) < \sum_{k=1}^{\bar{n}} \|g(\bar{t}_k) - g(\bar{t}_{k-1})\| + \delta/L_f$. So

$$\begin{aligned}
\sum_{k=1}^{n} \|f \circ g(\gamma(t_k)) - f \circ g(\gamma(t_{k-1}))\| &\le L_f \sum_{k=1}^{n} \|g(t_k) - g(t_{k-1})\| \\
&\le L_f \rho(g(x), g(y)) \\
&< L_f \left(\sum_{k=1}^{\bar{n}} \|g(\bar{t}_k) - g(\bar{t}_{k-1})\| + \delta/L_f \right) \\
&< L_f \left(L_g \sum_{k=1}^{\bar{n}} \|\bar{t}_k - \bar{t}_{k-1}\| + \delta/L_f \right) \\
&\le L_f L_g \ell(\gamma) + \delta\,.
\end{aligned}$$

Since $\delta > 0$ was arbitrary, we get the result. □

3.2 The form of the set of discontinuities

We are going to generalize the idea of transforming a discontinuous drift into a Lipschitz one to general dimensions.

For this we assume that the drift coefficient μ is piecewise Lipschitz in the sense of Definition 3.4, that is, there exists a hypersurface Θ with finitely many components such that $\mu|_{\mathbb{R}^d \setminus \Theta}$ is intrinsic Lipschitz. The assumption on the drift that will make our method work therefore encompasses assumptions on Θ.

Assumption 3.1. The drift coefficient μ is a piecewise Lipschitz function $\mathbb{R}^d \longrightarrow \mathbb{R}^d$. Its exceptional set Θ is a C^3 hypersurface.

A consequence of Assumption 3.1 is that locally there exists a C^2 orthonormal vector, that is, for every sufficiently small open and connected $B \subseteq \Theta$ there exists an orthonormal vector on B, i.e., a C^2-function $n : B \longrightarrow \mathbb{R}^d$ such that for all $\xi \in B$ the vector $n(\xi)$ is orthogonal to the tangent space of Θ in ξ and $\|n(\xi)\| = 1$. It is well-known, that there are in general two possible choices for n and that one can take $B = \Theta$ only if Θ is orientable. But given n on B, the only other orthonormal vector is $-n$.

Assumption 3.2. There exists a constant K such that $\|n'(\xi)\| \leq K$ for every $\xi \in \Theta$ and every orthonormal vector n on Θ.

Define the distance $d(x, \Theta)$ between a point x and the hypersurface Θ in the usual way, $d(x, \Theta) := \inf\{\|x - y\| : y \in \Theta\}$. For every $\varepsilon > 0$ we define $\Theta^\varepsilon := \{x \in \mathbb{R}^d : d(x, \Theta) < \varepsilon\}$.

Assumption 3.3. There exists $\varepsilon_0 > 0$ such that Θ^{ε_0} has the *unique closest point property*, i.e., for every $x \in \mathbb{R}^d$ with $d(x, \Theta) < \varepsilon_0$ there is a unique $p \in \Theta$ with $d(x, \Theta) = \|x - p\|$.

A set possessing the property described in Assumption 3.3 is called a *set of finite reach*. This and the notion of *unique closest point property* has been defined in [11].

Due to Assumption 3.3 there exists an $\varepsilon_0 > 0$ for which we may define a mapping $p : \Theta^{\varepsilon_0} \longrightarrow \Theta$ assigning to each x the point $p(x)$ in Θ closest to x.

Note that one can find examples of hypersurfaces which satisfy Assumption 3.2, but not Assumption 3.3, see Figure 1.

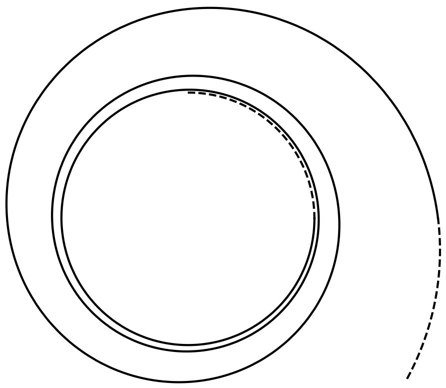

Figure 1: A hypersurface in $\mathbb{R}^2$ that satisfies Assumption 3.2, but not Assumption 3.3.

Lemma 3.10. *If Θ satisfies Assumptions 3.2 and 3.3, then item 3 of Lemma 3.6 is satisfied.*

The rather technical proof of this lemma can be found in the appendix. Note that for many examples, like a (hyper-)sphere or hyperplane, item 3 of Lemma 3.6 is obviously satisfied. So in these cases there is no need to resort to Lemma 3.10. However, it is an interesting fact that this condition is automatically satisfied under our assumptions on Θ.

3.3 Construction of the transform G

As before, we construct a transform G with the property that the SDE for $G(X)$ has Lipschitz coefficients.

For this to be well-defined, we make the following assumptions:

Assumption 3.4. There is a constant $c_0 > 0$ such that $\|\sigma(\xi)^\top n(\xi)\| \geq c_0$ for all $\xi \in \Theta$.

Remark 3.11. Assumption 3.4 is a *non-parallelity condition*, meaning that for all $\xi \in \Theta$, $\sigma(\xi)$ must not be parallel to Θ, i.e., there exists $x \in \mathbb{R}^d$ such that $\sigma(\xi)x$ is not in the tangent space of Θ in ξ.

For defining the transform, we first switch to a local setting. Suppose $\tilde{x} \in \mathbb{R}^d$ is close to Θ, i.e., $d(\tilde{x}, \Theta) < \varepsilon_0$. Let $B \subseteq \Theta$ be an open environment of $p(x)$ in Θ and n an orthonormal vector. It follows that the set

$$U = \{y_1 n(\xi)) + \xi : y_1 \in (-\varepsilon_0, \varepsilon_0), \xi \in B\}$$

is an open environment of $\tilde{x}$, and every point $x \in U$ can be uniquely represented in the form $x = y_1 n(\xi) + \xi$, $y_1 \in (-\varepsilon_0, \varepsilon_0)$, $\xi \in B$.

We are now ready to locally define the transform $G : U \longrightarrow \mathbb{R}^d$ by

$$G(x) = x + \tilde{\phi}(x)\alpha(p(x)) , \tag{6}$$

where $\tilde{\phi}(x) = (x - p(x)) \cdot n(p(x))\|x - p(x)\|\phi\left(\frac{\|x-p(x)\|}{c}\right)$ and where

$$\alpha(\xi) := \lim_{h\to 0} \frac{\mu(\xi - hn(\xi)) - \mu(\xi + hn(\xi))}{2n(\xi)^\top \sigma(\xi)\sigma(\xi)^\top n(\xi)} , \qquad \xi \in B . \tag{7}$$

One important point to note is the following proposition.

Proposition 3.12. *The value of the function G does not depend on the choice of the orthonormal vector.*

Proof. Both $\alpha(p(x))$ and $\tilde{\phi}(x)$ depend on the parametrization only through the direction of the normal vector $n(p(x))$. But from the definitions of $\tilde{\phi}$ and α we see that if $n(p(x))$ is replaced by $-n(p(x))$, then $\tilde{\phi}(x)$ and $\alpha(p(x))$ both change sign. Therefore, $\tilde{\phi}(x)\alpha(p(x))$ does not depend on the particular choice of the orthonormal vector. □

The only reason why we defined G locally at first was that for a non-orientable hypersurface we do not have, by definition, a global orthonormal vector. However, since the value of the locally defined function G does not depend on the particular choice of the orthonormal vector, we can use the same equations (6) and (7) for defining G globally on Θ^{ε_0}. That is, the function $G : \mathbb{R}^d \longrightarrow \mathbb{R}^d$,

$$G(x) = \begin{cases} x + \tilde{\phi}(x)\alpha(p(x)) & x \in \Theta^{\varepsilon_0} , \\ x & x \in \mathbb{R}^d \backslash \Theta^{\varepsilon_0} \end{cases}$$

is well-defined. Note further that, if we require $c \le \varepsilon_0$, then from $d(x, \Theta) > \varepsilon_0$ it follows $d(x, \Theta) > c$ and therefore $\phi(\frac{\|x-p(x)\|}{c}) = 0$ with a C^2 smooth paste to 0 in all points x satisfying $d(x, \theta) = c$.

3.4 Properties of G

We need to prove the following:

1. c can be chosen in a way such that G is a diffeomorphism $\mathbb{R}^d \longrightarrow \mathbb{R}^d$;
2. Itô's formula holds for G^{-1};
3. the SDE for $G(X)$ has Lipschitz coefficients.

Assumption 3.5. There is a constant a such that every locally defined function α as defined in (7) is C^3 and all derivatives up to order 3 are bounded by a.

Theorem 3.13. *Let Assumptions 3.1–3.5 be satisfied. If the constant $c > 0$ appearing in the definition of $\tilde{\phi}$ is sufficiently small, then G is a diffeomorphism $\mathbb{R}^d \longrightarrow \mathbb{R}^d$.*

For proving Theorem 3.13 we first need to prove two technical lemmas. For every $\xi \in \Theta$, denote by $\tau(\xi)$ the tangent space of Θ in ξ.

Lemma 3.14. *For $\xi \in \Theta$, n' is a linear mapping from $\tau(\xi)$ into $\tau(\xi)$.*

Proof. n' is by definition a linear mapping $\tau(\xi) \longrightarrow \mathbb{R}^d$. Furthermore, we have $\|n\| = 1$, so that for any curve γ in Θ

$$0 = \frac{d}{dt}\|n(\gamma(t))\|^2 = 2n(\gamma(t)) \cdot \left(n'(\gamma(t))\gamma'(t)\right) .$$

If $b \in \tau(\xi)$, we can find a curve γ in Θ such that $\gamma(0) = \xi$ and $\gamma'(0) = b$. Thus, $n(\xi) \cdot (n'(\xi)b) = 0$, i.e., $n'(\xi)b \in \tau(\xi)$. □

Remark 3.15. Under Assumptions 3.2 and 3.3 we may choose $0 < \varepsilon < \varepsilon_0$ such that, whenever $y_1 \in \mathbb{R}$ with $|y_1| < \varepsilon$, then $\mathrm{id}_{\tau(\xi)} + y_1 n'(\xi)$ is invertible .

Indeed, let $\varepsilon = \frac{\varepsilon_0}{\varkappa K}$ for some fixed $\varkappa > 1$. Then for $|y_1| < \varepsilon$ we have $\|y_1 n'(\xi)\| = |y_1|\|n'(\xi)\| < \frac{1}{\varkappa} < 1$, such that $\mathrm{id}_{\tau(\xi)} + y_1 n'(\xi)$ is invertible by the subsequent well-known Lemma 3.16.

Lemma 3.16. *Let $\mathcal{A}$ be a linear operator on a subspace $V \subseteq \mathbb{R}^d$ and let $\mathcal{A}$ have (operator) norm smaller than* 1*.*

Then $\mathrm{id}_V + \mathcal{A}$ *is invertible and* $\|(\mathrm{id}_V + \mathcal{A})^{-1}\| \leq (1 - \|\mathcal{A}\|)^{-1}$.

Proof. Invertibility is seen by considering the Neumann series $\mathcal{B} = \sum_{k=0}^{\infty}(-\mathcal{A})^k$ which converges in operator norm and satisfies $\|\mathcal{B}\| \leq (1 - \|\mathcal{A}\|)^{-1}$. Then

$$(\mathrm{id}_V + \mathcal{A})\mathcal{B} = \sum_{k=0}^{\infty}(-\mathcal{A})^k - \sum_{k=0}^{\infty}(-\mathcal{A})(-\mathcal{A})^k = \sum_{k=0}^{\infty}(-\mathcal{A})^k - \sum_{k=1}^{\infty}(-\mathcal{A})^k = \mathrm{id}_V .$$

□

From now on we fix some $\varkappa > 1$ and set $\varepsilon = \frac{\varepsilon_0}{\varkappa K}$, where K is the constant from Assumption 3.2.

Proof of Theorem 3.13. Let $0 < c < \varepsilon$.

For $\tilde{x} \notin \Theta^c$, differentiability of G in $\tilde{x}$ is obvious.

For $\tilde{x} \in \Theta^c$ choose an open subset B of Θ (as before) and an orthonormal vector n such that $U \subset \mathbb{R}^d$ is an open set with $U \cap \Theta = B$ and every $x \in U$ can uniquely be written in the form $x = y_1 n(\xi) + \xi$ with $\xi = p(x)$. Θ can be parametrized locally by a one-one mapping $\psi : R \to \mathbb{R}^d$, where $R \subseteq \mathbb{R}^{d-1}$ is an open rectangle in $\mathbb{R}^{d-1}$, and there is a point $(\tilde{y}_2, \dots, \tilde{y}_d) \in R$ such that $\psi(\tilde{y}_2, \dots, \tilde{y}_d) = p(\tilde{x})$. By making R and/or B smaller, if necessary, we may w.l.o.g. assume that $B = \psi(R)$.

Thus, we have a bijective mapping $\mathscr{T} : (-\varepsilon, \varepsilon) \times R \longrightarrow U$,

$$\mathscr{T}(y_1, \dots, y_d) := y_1 n(\psi(y_2, \dots, y_d)) + \psi(y_2, \dots, y_d) , \quad y \in (-\varepsilon, \varepsilon) \times R .$$

Note that $p(\mathscr{T}(y)) = \psi(y_2, \dots, y_d)$ for all $y \in (-\varepsilon, \varepsilon) \times R$.

We have

$$\begin{aligned} G \circ \mathscr{T}(y) &= y_1 n(\psi(y_2, \dots, y_d)) + \psi(y_2, \dots, y_d) + y_1|y_1|\phi\left(\frac{|y_1|}{c}\right)\alpha(\psi(y_2, \dots, y_d)) \\ &= y_1 n(\psi(y_2, \dots, y_d)) + \psi(y_2, \dots, y_d) + \bar{\phi}(y_1)\alpha(\psi(y_2, \dots, y_d)) , \end{aligned}$$

where $\bar{\phi} = y|y|\phi\left(\frac{y}{c}\right)$, and thus

$$\frac{\partial(G \circ \mathscr{T})}{\partial y_1}(y) = n(\psi(y_2, \dots, y_d)) + \bar{\phi}'(y_1)\alpha(\psi(y_2, \dots, y_d)) ,$$

$$\text{and} \quad \frac{\partial(G \circ \mathscr{T})}{\partial y_j}(y) = y_1 \frac{\partial(n \circ \psi)}{\partial y_j}(y_2, \dots, y_d) + \frac{\partial \psi}{\partial y_j}(y_2, \dots, y_d) + \bar{\phi}(y_1)\frac{\partial(\alpha \circ \psi)}{\partial y_j}(y_2, \dots, y_d) .$$

Now note that

$$\frac{\partial(G\circ\mathscr{T})}{\partial y_1}(y) = G'(\mathscr{T}(y))\frac{\partial\mathscr{T}}{\partial y_1}(y) = G'(\mathscr{T}(y))n(\psi(y_2,\dots,y_d))\,,$$

$$\text{and}\quad \frac{\partial(G\circ\mathscr{T})}{\partial y_j}(y) = G'(\mathscr{T}(y))\frac{\partial\mathscr{T}}{\partial y_j}(y) = G'(\mathscr{T}(y))\left(y_1\frac{\partial(n\circ\psi)}{\partial y_j}(y_2,\dots,y_d) + \frac{\partial\psi}{\partial y_j}(y_2,\dots,y_d)\right)$$

for all $j\neq 1$. Further,

$$\frac{\partial(n\circ\psi)}{\partial y_j}(y_2,\dots,y_d) = n'(\psi(y_2,\dots,y_d))\frac{\partial\psi}{\partial y_j}(y_2,\dots,y_d)\,,$$

$$\text{and}\quad \frac{\partial(\alpha\circ\psi)}{\partial y_j}(y_2,\dots,y_d) = \alpha'(\psi(y_2,\dots,y_d))\frac{\partial\psi}{\partial y_j}(y_2,\dots,y_d)\,.$$

Recall that for any $\xi\in\Theta$, we have that $n'(\xi)$ and $\alpha'(\xi)$ are linear mappings from the tangent space of Θ in ξ into the $\mathbb{R}^d$. For $\xi=\psi(y_2,\dots,y_d)$ it then follows that

$$G'(\mathscr{T}(y))\left(\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)\right)\frac{\partial\psi}{\partial y_j}(y_2,\dots,y_d) = \left(\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)+\bar\phi(y_1)\alpha'(\xi)\right)\frac{\partial\psi}{\partial y_j}(y_2,\dots,y_d)\,.$$

Since this equation holds for all $\frac{\partial\psi}{\partial y_j}$, $j=2,\dots,d$, it also holds for every vector b in the tangent space, i.e.,

$$G'(\mathscr{T}(y))\left(\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)\right)b = \left(\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)+\bar\phi(y_1)\alpha'(\xi)\right)b\,.$$

For $|y_1|\le\varepsilon$, the mapping $\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)$ is invertible by the argument from Remark 3.15. Denote the inverse of $\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)$ by $\mathcal{I}_\xi(y)$.

Then for any $b\in\tau(\xi)$ we can write $b=\left(\mathrm{id}_{\tau(\xi)}+y_1n'(\xi)\right)b_1$ with $b_1=\mathcal{I}_\xi(y)b$ and therefore

$$G'(\mathscr{T}(y))b = b+\bar\phi(y_1)\alpha'(\xi)\mathcal{I}_\xi(y)b\,.$$

For a general vector $b\in\mathbb{R}^d$ we have that $(b\cdot n)n = nn^\top b$ is orthogonal to the tangent space and $(\mathrm{id}_{\mathbb{R}^d}-nn^\top)b$ is in the tangent space.

We abbreviate $G'=G'(\tilde x)$, $p=p(\tilde x)$, $d=\|\tilde x-p(\tilde x)\|$, $n=n(p(\tilde x))$, $n'=n'(p(\tilde x))$, $\mathcal{I}_\xi=\mathcal{I}_\xi(\mathscr{T}^{-1}(\tilde x))$. Then we have for $b\in\mathbb{R}^d$

$$\begin{aligned}G'b &= G'\Big((b\cdot n)n+(b-(b\cdot n)n)\Big)\\ &= (b\cdot n)G'n+G'(b-(b\cdot n)n)\\ &= (b\cdot n)(n+\bar\phi'(d)\alpha(p))+(b-(b\cdot n)n)+\bar\phi(d)\alpha'(p)\mathcal{I}_\xi(b-(b\cdot n)n)\\ &= b+\bar\phi'(d)\alpha(p)n^\top b+\bar\phi(d)\alpha'(p)\mathcal{I}_\xi(\mathrm{id}_{\mathbb{R}^d}-nn^\top)b\,.\end{aligned}$$

Therefore,

$$G' = \mathrm{id}_{\mathbb{R}^d}+\bar\phi'(d)\alpha(p)n^\top+\bar\phi(d)\alpha'(p)\mathcal{I}_\xi(\mathrm{id}_{\mathbb{R}^d}-nn^\top)\,,$$

or, more explicitly,

$$\begin{aligned}G'(\tilde x) =&\mathrm{id}_{\mathbb{R}^d}+\bar\phi'(\|\tilde x-p(\tilde x)\|)\alpha(p(\tilde x))n(p(\tilde x))^\top\\ &+\bar\phi(\|\tilde x-p(\tilde x)\|)\alpha'(p(\tilde x))\mathcal{I}_\xi(\mathscr{T}^{-1}(\tilde x))(\mathrm{id}_{\mathbb{R}^d}-n(p(\tilde x))n(p(\tilde x))^\top)\,.\end{aligned}\tag{8}$$

In order to apply Hadamard's global inverse function theorem [15, Theorem 2.2], and thus to show that G is a diffeomorphism $\mathbb{R}^d\longrightarrow\mathbb{R}^d$, we need to show that G is C^1, $G'(x)>0$ is invertible for all $x\in\mathbb{R}^d$, and $\lim_{\|x\|\to\infty}\|G(x)\|=\infty$.

We have already proven differentiability of G in $\tilde x$. If c is sufficiently small, $G'(\tilde x)$ is invertible, since $\bar\phi'$ and $\bar\phi$ are uniformly bounded with a bound that tends to 0 for $c\to 0$. For c small enough it is therefore guaranteed that $G'(\tilde{x})$ is close to the identity and therefore invertible by Lemma 3.16. We show in the separate Lemma 3.17 that $c > 0$ can be chosen uniformly for all $\tilde{x}$ such that $G'(\tilde{x})$ is invertible.

Since $G(x) = x + \bar{\phi}(x)\alpha(x)$ and both $\bar{\phi}$ and α are bounded by the definition of $\bar{\phi}$ and Assumption 3.5, respectively, we also have the third requirement of Hadamard's global inverse function theorem. G is therefore a diffeomorphism. $\square$

We will see that c can always be chosen sufficiently small in the proof of Theorem 3.13.

Lemma 3.17. *Let*

$$c < -\frac{256\,(\varkappa-1)\,d\,|\alpha_i(p(x))|}{27\,\varkappa\,(d-1)\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right|} + \sqrt{\frac{256\,(\varkappa-1)}{27\,\varkappa\,(d-1)\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right|}\left(\frac{256\,(\varkappa-1)\,d^2\,|\alpha_i(p(x))|^2}{27\,\varkappa\,(d-1)\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right|} + \frac{1}{d}\right)}$$

for all $i, j = 1, \dots, d$.

With this choice of c *we have that* G' *is invertible on the whole of* $\mathbb{R}^d$.

Proof. In the proof of Theorem 3.13 we have seen that

$$\begin{aligned} G'(x) &= \mathrm{id}_{\mathbb{R}^d} + \bar{\phi}'(\|x-p(x)\|)\alpha(p(x))n(p(x))^\top \\ &\quad + \bar{\phi}(\|x-p(x)\|)\alpha'(p(x))\mathcal{I}_\xi(\mathscr{T}^{-1}(x))(\mathrm{id}_{\mathbb{R}^d} - n(p(x))n(p(x))^\top) =: 1 + \mathcal{A}\,, \end{aligned}$$

where $\mathscr{T}$ is as before. We begin by estimating the operator norm of $\mathcal{A}$.

$$\begin{aligned} \|\mathcal{A}\| &\le \sum_{i=1}^d \|\bar{\phi}'(\|x-p(x)\|)\| \sum_{j=1}^d |\alpha_i(p(x))n_j(p(x))| \\ &\quad + d(d-1)\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right| |\bar{\phi}(\|x-p(x)\|)| \|\mathcal{I}_\xi\| \|\mathrm{id}_{\mathbb{R}^d} - n(p(x))n(p(x))^\top\| \\ &\le \max_{1\le i,j\le d} d^2 |\alpha_i(p(x))| \|\bar{\phi}'\| \|x-p(x)\| \\ &\quad + d(d-1)\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right| |\bar{\phi}(\|x-p(x)\|)| \|\mathcal{I}_\xi\| \|\mathrm{id}_{\mathbb{R}^d} - n(p(x))n(p(x))^\top\| \\ &\le \max_{1\le i,j\le d} 2\,c\,d^2|\alpha_i(p(x))| + \frac{27c^2}{256} d(d-1)\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right| \frac{1}{1+|y_1|\|n'\|}\,, \end{aligned}$$

where we used that $\|x-p(x)\| \le c$ and $\|\bar{\phi}'\| \le 2$ for $x \in \Theta^c$, $|\bar{\phi}(\|x-p(x)\|)|$ attains its maximum in $\frac{27c^2}{256}$, and $\|\mathrm{id}_{\mathbb{R}^d} - n(p(x))n(p(x))^\top\| \le 1$. Furthermore $\|\mathcal{I}_\xi\| \le \frac{1}{1+|y_1|\|n'\|}$, since $\|y_1 n'\| < \frac{1}{\varkappa}$ by Lemma 3.16 and Remark 3.15, respectively.

Then $\frac{1}{1+|y_1|\|n'\|} \le \frac{\varkappa}{\varkappa-1}$.

Hence, to complete the proof we have to solve the quadratic inequality

$$\frac{27\,c^2\,\varkappa\,d\,(d-1)}{256\,(\varkappa-1)}\left|\frac{\partial \alpha_i(p(x))}{\partial x_j}\right| + 2\,c\,d^2|\alpha_i(p(x))| < 1$$

in c to get the second upper bound for c, such that $G'(x) > 0$ for $x \in \Theta^c$ by Lemma 3.16. For $x \in \mathbb{R}^d \backslash \Theta^c$, $G'(x) = \mathrm{id}_{\mathbb{R}^d}$. $\square$

W.l.o.g. we always choose c like in Lemma 3.17.

We proceed with proving that, although $G \notin C^2$, Itô's formula holds for G and G^{-1}.

Theorem 3.18. *Let Assumptions 3.1–3.5 be satisfied.*

Then Itô's formula holds for G *and* G^{-1}.

Proof. If $x \in \mathbb{R}^d \backslash \Theta$, then since $G, G^{-1} \in C^2$ on $\mathbb{R}^d \backslash \Theta$, Itô's formula holds for G and G^{-1} until the first time X hits Θ. So the only interesting case is $x \in \Theta$.

For this, there exists an open rectangle $R \in \mathbb{R}^{d-1}$ and a local parametrization $\psi : R \longrightarrow \mathbb{R}^d$ of Θ. Let $B = \psi(R)$. Moreover,

$$U = \{y_1 n(\psi(y_2, \ldots, y_d)) + \psi(y_2, \ldots, y_d) : y_1 \in (-\varepsilon, \varepsilon), (y_2, \ldots, y_d) \in R\}.$$

Let $\mathscr{T} : (-\varepsilon, \varepsilon) \times R \longrightarrow U$ be defined as in the proof of Theorem 3.13. Note that $\mathscr{T} \in C^2$, because Θ is C^3 by Assumption 3.1, so Itô's formula holds for $\mathscr{T}$. $\mathscr{T}$ is locally invertible with $\det \mathscr{T}' \neq 0$, so $\mathscr{T}^{-1} \in C^2$ as well. So if we can show that Itô's formula holds for $G \circ \mathscr{T}$, then it also holds for $G = G \circ \mathscr{T} \circ \mathscr{T}^{-1}$.

$G \circ \mathscr{T}$ fits the assumptions of [14, Theorem 2.9] (we get boundedness of the derivatives by localizing to a bounded domain), so Itô's formula holds for $G \circ \mathscr{T}$, and therefore also for G.

$\tilde{G} = \mathscr{T}^{-1} \circ G \circ \mathscr{T}$ is a function with continuous first and second derivatives, with the sole exception of $\frac{\partial^2 \tilde{G}}{\partial y_1^2}$, which is bounded, but may be discontinuous for $y_1 = 0$. Since $\det \tilde{G}' \neq 0$ on an environment of x, this property transfers to the inverse, which is $\tilde{G}^{-1} = \mathscr{T}^{-1} \circ G^{-1} \circ \mathscr{T}$. Thus, again by [14, Theorem 2.9], Itô's formula holds for $\tilde{G}^{-1}$, and a fortiori for G^{-1}. □

Now we are ready to show that the coefficients of the transformed SDE for $G(X)$ are Lipschitz.

Assumption 3.6. We assume the following for μ and σ:

1. the diffusion coefficient σ is Lipschitz;
2. μ and σ are bounded on Θ^ε.

Theorem 3.19. *Let Assumptions 3.1–3.6 be satisfied.*
Then the SDE for $G(X)$ has Lipschitz coefficients.

Proof. We first show that the drift of $G(X)$ is continuous in Θ. Let B, R, ψ, and $\mathscr{T}$ be defined as in the proof of Theorem 3.13. Suppose now, we have a locally defined process X in U. Then there exists a locally defined process Y in $(-\varepsilon, \varepsilon) \times R$ with

$$X = Y_1 n(\psi(Y_2, \ldots, Y_d)) + \psi(Y_2, \ldots, Y_d),$$

i.e., $X = \mathscr{T}(Y)$.

If Y is a locally defined solution to $dY = \nu(Y)dt + \omega(Y)dW$, then by Itô's formula

$$dX = \mathscr{T}'(Y)\nu(Y)dt + \mathscr{T}'(Y)\omega(Y)dW + \frac{1}{2}\operatorname{tr}(\omega^\top(Y)\mathscr{T}''(Y)\omega(Y))dt,$$

where $\mathscr{T}'$ and $\mathscr{T}''$ denote the Jacobian and the Hessian of $\mathscr{T}$, respectively, and tr denotes the trace of a matrix. We want $\mathscr{T}'\omega = \sigma$, or more precisely $\mathscr{T}'(Y)\omega(Y) = \sigma(\mathscr{T}(Y))$, i.e., $\omega = (\mathscr{T}')^{-1}\sigma$. For brevity write $\mathscr{S} = \mathscr{T}^{-1}$. Now

$$\begin{aligned}(\omega\omega^\top)_{1,1} &= \omega_{1,1}^2 + \cdots + \omega_{1,d}^2 = e_1^\top \omega\omega^\top e_1 \\ &= e_1^\top \left(\mathscr{S}'\sigma\sigma^\top \left(\mathscr{S}'\right)^\top\right) e_1 .\end{aligned}$$

We show that $(\mathscr{S}')^\top e_1 = n$. It is not hard to see that the Jacobian $\mathscr{T}'$ of $\mathscr{T}$ in a point $\xi \in \Theta$ is given by

$$\mathscr{T}' = \left(n, \frac{\partial \psi}{\partial y_2}, \ldots, \frac{\partial \psi}{\partial y_d}\right),$$

such that

$$e_1^\top (\mathscr{T}')^{-1} = e_1^\top \left((\mathscr{T}')^{-1}\right) = n^\top \iff e_1^\top = n^\top \mathscr{T}' = n^\top \left(n, \frac{\partial \psi}{\partial y_2}, \dots, \frac{\partial \psi}{\partial y_d}\right) = (\|n\|^2, 0, \dots, 0) = e_1^\top \, .$$

Therefore we have $\omega_{1,1}^2 + \cdots + \omega_{1,d}^2 = n^\top \sigma \sigma^\top n$ on Θ.

The drift coefficient ν of the SDE for Y has only discontinuities in the set $\{y \in \mathbb{R}^d : y_1 = 0\}$. Further,

$$dY = d(\mathscr{S}(X)) = \mathscr{S}'(X)\mu(X)dt + \mathscr{S}'(X)\sigma(X)dW + \frac{1}{2}\operatorname{tr}\left(\sigma^\top(X)\mathscr{S}''(X)\sigma(X)\right)dt \, ,$$

i.e., $\nu(y) = \mathscr{S}'(\mathscr{T}(y))\mu(\mathscr{T}(y)) + \frac{1}{2}\operatorname{tr}\left(\sigma^\top(\mathscr{T}(y))\mathscr{S}''(\mathscr{T}(y))\sigma(\mathscr{T}(y))\right)$. The second term is continuous, so that

$$\begin{aligned}
&\lim_{h\to 0}(\nu(-h, y_2, \dots, y_d) - \nu(h, y_2, \dots, y_d)) \\
&= \mathscr{S}'(\mathscr{T}(0, y_2, \dots, y_d)) \lim_{h\to 0} \left(\mu(\mathscr{T}(-h, y_2, \dots, y_d)) - \mu(\mathscr{T}(h, y_2, \dots, y_d))\right) \\
&= \mathscr{S}'(\mathscr{T}(0, y_2, \dots, y_d)) \lim_{h\to 0} \Big(\mu\big(\mathscr{T}(0, y_2, \dots, y_d) - hn(\mathscr{T}(0, y_2, \dots, y_d))\big) \qquad (9)\\
&\qquad\qquad - \mu\big(\mathscr{T}(0, y_2, \dots, y_d) + hn(\mathscr{T}(0, y_2, \dots, y_d))\big)\Big) \\
&= \mathscr{S}'(\mathscr{T}(0, y_2, \dots, y_d)) 2\alpha(\mathscr{T}(0, y_2, \dots, y_d))(n^\top \sigma\sigma^\top n)(\mathscr{T}(0, y_2, \dots, y_d)) \\
&= \mathscr{S}'(\mathscr{T}(0, y_2, \dots, y_d)) 2\alpha(\mathscr{T}(0, y_2, \dots, y_d))(\omega\omega^\top)_{11}(0, y_2, \dots, y_d) \, . \qquad (10)
\end{aligned}$$

Consider

$$\begin{aligned}
(G \circ \mathscr{T})(y) &= \mathscr{T}(y) + \tilde{\phi}(\mathscr{T}(y))\alpha(p(\mathscr{T}(y))) \\
&= y_1 n(\mathscr{T}(0, y_2, \dots, y_d)) + \mathscr{T}(0, y_2, \dots, y_d) + y_1|y_1|\phi\left(\frac{y_1}{c}\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right) \, ,
\end{aligned}$$

and

$$(\mathscr{S} \circ G \circ \mathscr{T})(y) = \mathscr{S}\left(y_1 n(\mathscr{T}(0, y_2, \dots, y_d)) + \mathscr{T}(0, y_2, \dots, y_d) + y_1|y_1|\phi\left(\frac{y_1}{c}\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right)\right) \, .$$

Differentiation yields

$$\begin{aligned}
\frac{\partial}{\partial y_1}(\mathscr{S} \circ G \circ \mathscr{T})(y) &= \mathscr{S}'\left((G \circ \mathscr{T})(y)\right) \frac{\partial}{\partial y_1}\Big(y_1 n(\mathscr{T}(0, y_2, \dots, y_d)) + \mathscr{T}(0, y_2, \dots, y_d) \\
&\qquad + y_1|y_1|\phi\left(\frac{y_1}{c}\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right)\Big) \\
&= \mathscr{S}'\left((G \circ \mathscr{T})(y)\right)\Big(n(\mathscr{T}(0, y_2, \dots, y_d)) \\
&\qquad + \left(2|y_1|\phi\left(\frac{y_1}{c}\right) + y_1|y_1|\phi'\left(\frac{y_1}{c}\right)\frac{1}{c}\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right)\Big) \, .
\end{aligned}$$

We look at the second derivative w.r.t. y_1:

$$\frac{\partial^2}{\partial y_1^2}(\mathscr{S} \circ G \circ \mathscr{T})(y) = \text{something continuous} + \mathscr{S}'\left((G \circ \mathscr{T})(y)\right)\left(2\operatorname{sign}(y_1)\phi\left(\frac{y_1}{c}\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right)\right) \, .$$

Since $G(x) = x$ for $x \in \Theta$, we have that $G(\mathscr{T}(y)) = \mathscr{T}(y)$ for $y_1 = 0$, and thus

$$\begin{aligned}
&\lim_{h\to 0+}\left(\frac{\partial^2}{\partial y_1^2}(\mathscr{S} \circ G \circ \mathscr{T})(-h, y_2, \dots, y_d) - \frac{\partial^2}{\partial y_1^2}(\mathscr{S} \circ G \circ \mathscr{T})(h, y_2, \dots, y_d)\right) \\
&\quad = -4\mathscr{S}'\left((G \circ \mathscr{T})(0, y_2, \dots, y_d)\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right) \\
&\quad = -4\mathscr{S}'\left(\mathscr{T}(0, y_2, \dots, y_d)\right)\alpha\left(\mathscr{T}(0, y_2, \dots, y_d)\right) \, . \qquad (11)
\end{aligned}$$

Consider the drift coefficient of $(\mathscr{S}\circ G\circ \mathscr{T})_k(Y)$, which is

$$\tilde\nu_k(y) := \sum_{j=1}^d \frac{\partial}{\partial y_j}(\mathscr{S}\circ G\circ \mathscr{T})_k(y)\nu_j(y) + \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial y_i \partial y_j}(\mathscr{S}\circ G\circ \mathscr{T})_k(y)\sum_{l=1}^d \omega_{li}(y)\omega_{lj}(y)\,.$$

$(\mathscr{S}\circ G\circ \mathscr{T})'(0,y_2,\ldots,y_d) = \mathrm{id}_{\mathbb{R}^d}$, thus $\frac{\partial}{\partial y_j}(\mathscr{S}\circ G\circ \mathscr{T})_k(0,y_2,\ldots,y_d) = (e_k)_j$. Further, note that $\frac{\partial^2}{\partial y_i \partial y_j}(\mathscr{S}\circ G\circ \mathscr{T})_k$ is continuous for all pairs (i,j) except $(i,j)=(1,1)$.

Thus, using (10) and (11), we have

$$\begin{aligned}
&\lim_{h\to 0+}\left(\tilde\nu_k(-h,y_2,\ldots,y_d) - \tilde\nu_k(h,y_2,\ldots,y_d)\right)\\
&= \lim_{h\to 0+}\Big(\nu_k(-h,y_2,\ldots,y_d) + \frac{1}{2}\frac{\partial^2}{\partial y_1^2}(\mathscr{S}\circ G\circ \mathscr{T})_k(-h,y_2,\ldots,y_d)(\omega\omega^\top)_{11}(0,y_2,\ldots,y_d)\\
&\qquad - \nu_k(h,y_2,\ldots,y_d) - \frac{1}{2}\frac{\partial^2}{\partial y_1^2}(\mathscr{S}\circ G\circ \mathscr{T})_k(h,y_2,\ldots,y_d)(\omega\omega^\top)_{11}(0,y_2,\ldots,y_d)\Big)\\
&= \mathscr{S}'(\mathscr{T}(0,y_2,\ldots,y_d))2\alpha(\mathscr{T}(0,y_2,\ldots,y_d))(\omega\omega^\top)_{11}(0,y_2,\ldots,y_d)\\
&\quad - 2\mathscr{S}'\left(\mathscr{T}(0,y_2,\ldots,y_d)\right)\alpha\left(\mathscr{T}(0,y_2,\ldots,y_d)\right)(\omega\omega^\top)_{11}(0,y_2,\ldots,y_d) \;=\; 0\,.
\end{aligned}$$

Therefore $\tilde\nu$ is continuous on the whole of $\mathbb{R}^d$.

Now the drift coefficient of the SDE for the process $G(X)$ is continuous as well: $G(X) = \mathscr{T}\circ(\mathscr{S}\circ G\circ \mathscr{T})\circ\mathscr{S}(X)$ and compounding with $\mathscr{T}$ and $\mathscr{S}$ preserves continuity of the drift since $\mathscr{T},\mathscr{S}\in C^2$.

The k-th coordinate of the transformed drift $\tilde\mu$ has the form

$$\tilde\mu_k(z) = G_k'(G^{-1}(z))\mu(G^{-1}(z)) + \frac{1}{2}\operatorname{tr}\left(\sigma^\top(G^{-1}(z))G_k''(G^{-1}(z))\sigma(G^{-1}(z))\right)$$

and we have just seen that it is continuous in all $z\in\Theta$. It remains to show that $\tilde\mu$ is intrinsic Lipschitz on $\mathbb{R}^d\backslash\Theta$. For $z\in\mathbb{R}^d\backslash\Theta^c$ we have $\tilde\mu(z)=\mu(z)$. μ is intrinsic Lipschitz on $\mathbb{R}^d\backslash\Theta$, and therefore also on $z\in\mathbb{R}^d\backslash\Theta^c$.

On $\Theta^c\backslash\Theta$ we have that G' is differentiable with bounded derivative and is therefore intrinsic Lipschitz by Lemma 3.8. μ is intrinsic Lipschitz on $\mathbb{R}^d\backslash\Theta$ by Assumption 3.1 and μ is bounded on Θ^c by Assumption 3.6, item 2. Moreover, G^{-1} is Lipschitz on $\mathbb{R}^d$ and thus the mapping $x\mapsto G_k'(G^{-1}(z))\mu(G^{-1}(z))$ is intrinsic Lipschitz by Lemma 3.9.

In the same way we see that G'' is differentiable with bounded derivative on $\Theta^c\backslash\Theta$ and is therefore intrinsic Lipschitz by Lemma 3.8. σ is Lipschitz on $\mathbb{R}^d$ and therefore intrinsic Lipschitz on $\Theta^c\backslash\Theta$. Moreover, both G'' and σ are bounded on $\Theta^c\backslash\Theta$, thus $z\mapsto\frac{1}{2}\operatorname{tr}\left(\sigma^\top(G^{-1}(z))G_k''(G^{-1}(z))\sigma(G^{-1}(z))\right)$ is intrinsic Lipschitz by Lemma 3.9.

Now $\tilde\mu$ is intrinsic Lipschitz as a sum of intrinsic Lipschitz functions.

Altogether we have shown that $\tilde\mu$ is piecewise Lipschitz and continuous, and hence Lipschitz by Lemma 3.6 and Lemma 3.10.

The transformed diffusion coefficient is given by

$$\tilde\sigma(z) = G'(G^{-1}(z))\sigma(G^{-1}(z))\,.$$

Since G^{-1}, G' and σ are Lipschitz, the mappings $z\mapsto G'(G^{-1}(z))$ and $z\mapsto\sigma(G^{-1}(z))$ are Lipschitz. Moreover, they are both bounded on Θ^ε (and thus on Θ^c), such that their product is Lipschitz. □

3.5 Main results

Finally, we are ready to prove the two main results of this paper.

For this, define

$$dZ = dG(X) = \tilde{\mu}(Z)dt + \tilde{\sigma}(Z)dW\,, \qquad Z_0 = G(x)\,, \tag{12}$$

where $\tilde{\mu}$ and $\tilde{\sigma}$ are defined in the proof of Theorem 3.19.

Theorem 3.20. *Let Assumptions 3.1–3.6 be satisfied.*
Then the d-dimensional SDE (1) *has a unique global strong solution.*

Proof. Since by Theorem 3.19 SDE (12) has Lipschitz coefficients, it follows that it has a unique global strong solution for the initial value $G(x)$. Due to Theorem 3.13, the transformation G has a global inverse G^{-1}. Itô's formula holds for G^{-1} by Theorem 3.18. Applying Itô's formula to G^{-1}, we obtain that $G^{-1}(Z)$ satisfies

$$dX = \mu(X)dt + \sigma(X)dW\,, \quad X_0 = x\,.$$

Setting $X = G^{-1}(Z)$ closes the proof. □

For calculating the solution to the d-dimensional SDE (1), the same algorithm as for the one-dimensional case works, if applied using the transformations from the d-dimensional case. Let $Z_T^{(\delta)}$ be the Euler-Maruyama approximation of the solution to SDE (12) with step size smaller than $\delta > 0$.

Algorithm 3.21. Go through the following steps:

1. Set $Z_0^{(\delta)} = G(x)$.
2. Apply the Euler-Maruyama method to SDE (12) to obtain $Z_T^{(\delta)}$.
3. Set $\bar{X} = G^{-1}\left(Z_T^{(\delta)}\right)$.

Theorem 3.22. *Let Assumptions 3.1–3.6 be satisfied.*
Then Algorithm 3.21 converges with strong order $1/2$ to the solution X of the d-dimensional SDE (1).

Proof. We estimate the L^2-error of the approximation. For every $T > 0$ there is a constant C, such that

$$\mathbb{E}\left(\left\|X_T - \bar{X}_T\right\|^2\right) = \mathbb{E}\left(\left\|G^{-1}\left(Z_T\right) - G^{-1}\left(Z_T^{(\delta)}\right)\right\|^2\right) \le L_{G^{-1}}^2 \mathbb{E}\left(\left\|Z_T - Z_T^{(\delta)}\right\|^2\right) = L_{G^{-1}}^2 C\delta$$

for every sufficiently small step size δ, where $L_{G^{-1}}$ is the Lipschitz constant of G^{-1}. We used [9, Theorem 10.2.2] for the L^2-convergence of order $1/2$ of the Euler-Maruyama scheme for SDEs with Lipschitz coefficients. □

3.6 Compact set of discontinuities

To be able to prove our main results we had to make a number of assumptions on the coefficient functions μ and σ. At least one of those is indispensable for our method to work, that is, Assumption 3.1, which demands that μ is piecewise Lipschitz and that its set of discontinuities Θ is a C^3 hypersurface.

There are two more assumptions on Θ and several on the behaviour of the coefficients close to Θ. In this subsection we shall find out which assumptions are automatically satisfies in the case where Θ is compact.

In that case, Assumption 3.2 is certainly satisfied, since then n' is a continuous function on a compact set and is therefore bounded.

For compact Θ, Assumption 3.3 is automatically satisfied, too. This follows from a lemma in [4]:

Lemma 3.23. *Let $\Theta \subseteq \mathbb{R}^d$ be a compact C^k submanifold with $k \geq 2$.*

Then Θ has a neighbourhood $U = \Theta^\varepsilon$ with the unique closest point property, and the projection map $p : U \longrightarrow \Theta$ is C^{k-1}.

Assumption 3.4 prescribes a certain geometrical relation between Θ and directions of the diffusion coefficient. This will not be satisfied automatically only from making additional assumptions on Θ, of course. But for the case of constant Θ, Assumption 3.4 follows easily from weaker requirements on σ.

Proposition 3.24. *Let Θ be a compact C^2 hypersurface and let $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be Lipschitz.*

If $\sigma(\xi)^\top n(\xi) \neq 0$ for all $\xi \in \Theta$, then there exists a constant $c_0 > 0$ such that $\|\sigma^\top(\xi)n(\xi)\| \geq c_0$ for all $\xi \in \Theta$.

Proof. Let $B \subseteq \Theta$ be a bounded, open, and connected subset with the property that there exists an orthonormal vector n on B. Since $\sigma^\top n$ is continuous on the closure $\overline{B}$, there exists $c > 0$ such that $\|\sigma(\xi)^\top(\xi)\| \geq c$ for all $\xi \in B$.

By compactness, Θ can be covered by finitely many sets $B_1, \ldots, B_n$ with lower bounds $c_1, \ldots, c_n$ and we can take $c_0 := \min(c_1, \ldots, c_n)$ for the conclusion to hold. □

Note that $\sigma(\xi)^\top n(\xi) \neq 0$ also follows from $\det(\sigma(\xi)) \neq 0$. So in particular, regularity of σ implies Assumption 3.4 for compact Θ.

Finally, consider Assumption 3.5 which asserts boundedness of the first three derivatives of the locally defined function α on Θ. Similar to what we have done in the proof of Proposition 3.24, we can conclude boundedness of the derivatives from their continuity.

Assumption 3.6.2 is also automatically satisfied for compact Θ.

4 Numerical Example

In this section we present a concrete example. We compute the transform G as well as the coefficients $\tilde{\mu}, \tilde{\sigma}$ of the transformed SDE to which we apply the Euler-Maruyama scheme. Furthermore, we examine the quality of the approximation by considering the estimated L^2-error.

Let Θ be the unit circle in $\mathbb{R}^2$, i.e., the drift of our SDE is discontinuous only in $\Theta = \{x \in \mathbb{R}^2 : \|x\| = 1\}$. We want to solve the following SDE numerically:

$$\begin{pmatrix} dX \\ dY \end{pmatrix} = \mu(X,Y)dt + \sigma(X,Y)dW_t\,, \qquad \begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}, \tag{13}$$

where

$$\mu(x,y) = \begin{cases} (-x,-y)^\top, & x^2+y^2 > 1 \\ (x,0)^\top, & x^2+y^2 < 1\,, \end{cases}$$

$\sigma \equiv \mathrm{id}_{\mathbb{R}^2}$, and W is a two-dimensional standard Brownian motion.

We have that $p(x,y) = n(x,y) = (\sqrt{x^2+y^2})^{-1}(x,y)^\top$ yielding the transform

$$G(x,y) = \begin{cases} \left(1 + \frac{(\sqrt{x^2+y^2}-1)|\sqrt{x^2+y^2}-1|}{\sqrt{x^2+y^2}}\phi\left(\frac{|1-\sqrt{x^2+y^2}|}{c}\right)\right)\begin{pmatrix} x \\ y \end{pmatrix}, & (1+c)^2 > x^2+y^2 \geq 1 \\ \left(1 + \frac{(\sqrt{x^2+y^2}-1)|\sqrt{x^2+y^2}-1|}{2\sqrt{x^2+y^2}}\phi\left(\frac{|1-\sqrt{x^2+y^2}|}{c}\right)\right)\begin{pmatrix} x \\ y \end{pmatrix}, & (1-c)^2 < x^2+y^2 < 1\,, \end{cases}$$

and $G = \mathrm{id}_{\mathbb{R}^2}$, if $x^2 + y^2 \geq (1+c)^2$, or $x^2 + y^2 \leq (1-c)^2$, where we have chosen $c = 1/2$.

Then the drift of the transformed SDE is given by

$$\tilde{\mu}(G^{-1}(x,y)) = \begin{cases} \nabla G(x,y)(-x,-y)^\top + \frac{1}{2}\operatorname{tr}(G''(x,y)), & (1+c)^2 > x^2 + y^2 \geq 1 \\ \nabla G(x,y)(x,0)^\top + \frac{1}{2}\operatorname{tr}(G''(x,y)), & (1-c)^2 < x^2 + y^2 < 1\,, \end{cases}$$

and $\tilde{\mu}(x,y) = (-x,-y)^\top$, if $x^2 + y^2 \geq (1+c)^2$, and $\tilde{\mu}(x,y) = (x,0)^\top$, if $x^2 + y^2 \leq (1-c)^2$. Furthermore, $\tilde{\sigma}(G^{-1}(x,y)) = \nabla G(x,y)$. G^{-1} has to be evaluated numerically.

Figure 2 shows the deviation of the first component of G from identity. Figure 3 shows the first component of $\tilde{\mu}$, and Figure 4 shows $\tilde{\sigma}_{11}$. All other components look similar.

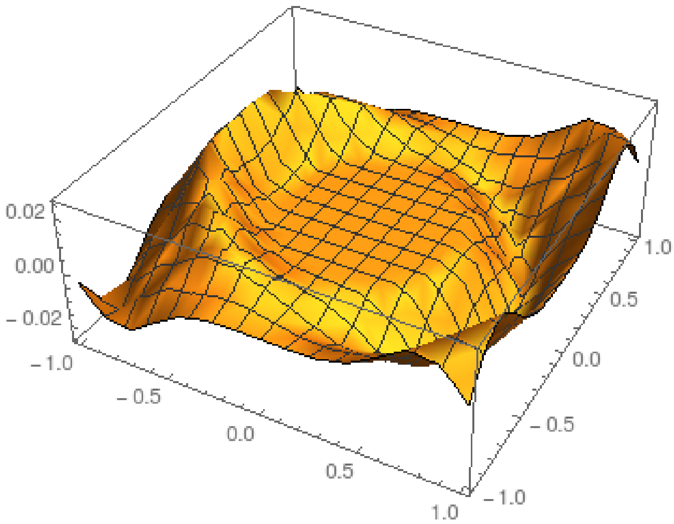

Figure 2: The function $(x,y) \mapsto G_1(x,y) - x$.

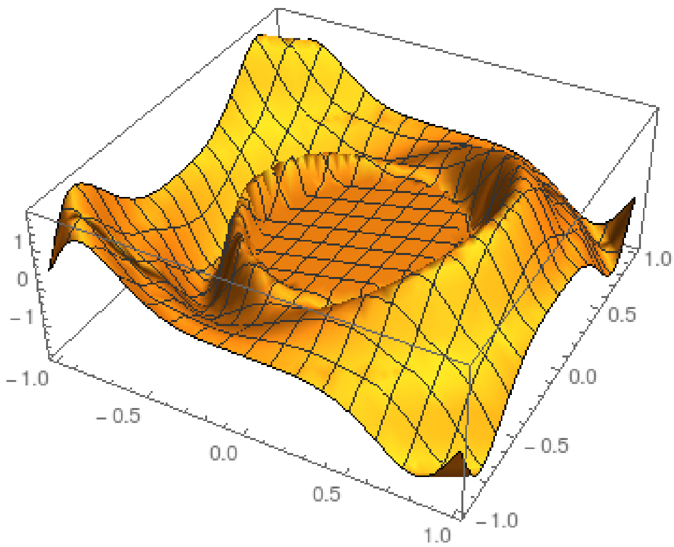

Figure 3: The function $\tilde{\mu}_1$.

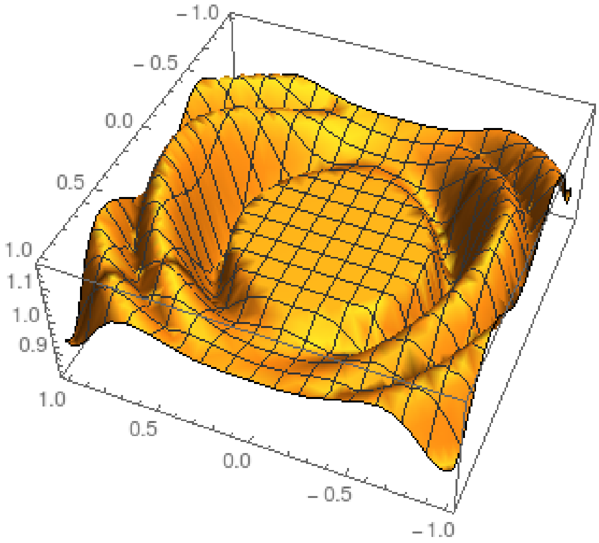

Figure 4: The function $\tilde{\sigma}_{11}$.

We apply Algorithm 3.21 to solve SDE (13). Figure 5 shows the estimated L^2-error of the approximation of our G-transformed Euler-Maruyama method (GM), compared to the Euler-Maruyama (EM) scheme:

$$err_k := \log_2 \left(d \sqrt{\hat{E}\left(\left\| X_T^{(k)} - X_T^{(k-1)} \right\|^2 \right)} \right)$$

plotted over $\log_2 \delta^{(k)}$, where $X_T^{(k)}$ is the numerical approximation with step size $\delta = \delta^{(k)}$, $\hat{E}$ is an estimator of the mean value using 1024 paths, and d is a normalizing so that $err_1 = \sqrt{1/2}$.

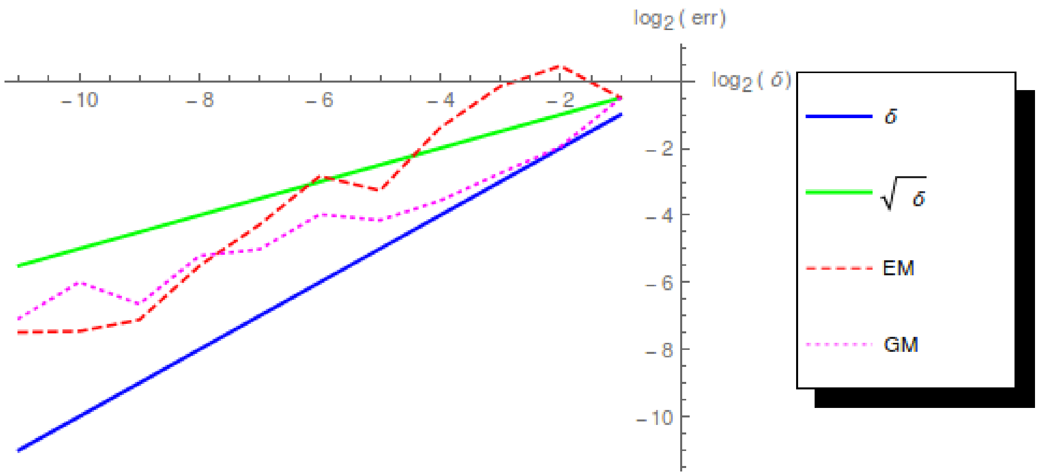

Figure 5: The estimated L^2-error.

We observe that for a relatively large step size δ our method (GM) gives a smaller error than crude Euler-Maruyama (EM), but for smaller step size, both methods seem to converge with the same order.

Note however, that it is an open problem to prove strong convergence of the Euler-Maruyama method.

Examples from stochastic control theory, where SDEs with discontinuous drift and degenerate diffusion coefficient appear are, e.g., [13, 17, 18, 19]. The SDEs appearing there can now be shown to have a unique global strong solution under conditions significantly weaker than known so far, and this solution can be approximated with a numerical method that converges with strong order $1/2$.

A Proof of Lemma 3.10

We now give a proof of Lemma 3.10, i.e., we prove the claim that a hypersurface that satisfies Assumptions 3.2 and 3.3 has the property that every line segment from x to y can be replaced by a continuous curve γ from x to y with $\ell(\gamma) < \|x-y\| + \eta$ where $\eta > 0$ is a given constant.

Let from now on $\varepsilon < \varepsilon_0$, where ε_0 is as in Assumption 3.3, so that in particular for every $x \in \mathbb{R}^d$ with $d(x, \Theta) \le \varepsilon$ there is a unique closest point $p(x)$ on Θ.

Denote by s the line segment from x to y and identify it with it's parameter representation $s(t) = x + t(y-x)\|y-x\|^{-1}$. Let $A := \{t \in [0, \|y-x\|] : s(t) \in \Theta\}$. For any set $S \subseteq \mathbb{R}$ denote by $H(S)$ the set of accumulation points of S.

Proposition A.1. *Let $t \in H(A)$. Then $n(s(t)) \perp s'(t)$.*

Proof. Suppose this was not the case, i.e., $n(s(t)) \cdot s'(t) \neq 0$. W.l.o.g. $n(s(t)) \cdot s'(t) = C > 0$. Let $(t_j)_{j\in\mathbb{N}}$ be a sequence in A with $t_j \neq t$, $\lim_j t_j = t$. W.l.o.g, $t_j > t$ for all j, or $t_j < t$ for all j.

By Assumption 3.3 we have $(B_\varepsilon(s(t) - \varepsilon n(s(t))) \cup B_\varepsilon(s(t) + \varepsilon n(s(t)))) \cap \Theta = \emptyset$, where $B_r(z)$ denotes the open ball with midpoint z and radius r.

Suppose $t_j > t$ for all j. Then

$$\begin{aligned}\|s(t) + \varepsilon n(s(t)) - s(t_j)\|^2 &= \|s(t) - s(t_j)\|^2 + 2\varepsilon(s(t) - s(t_j)) \cdot n(s(t)) + \varepsilon^2 \|n(s(t))\|^2 \\ &= \|s(t) - s(t_j)\|^2 + 2\varepsilon(t - t_j)s'(t) \cdot n(s(t)) + \varepsilon^2 \\ &= |t - t_j|^2 - 2\varepsilon|t - t_j|C + \varepsilon^2 \\ &= |t - t_j|(|t - t_j| - 2\varepsilon C) + \varepsilon^2\,,\end{aligned}$$

and the last expression is smaller than ε^2 for j large enough. Thus we have found a point ξ on Θ, namely $\xi = s(t_j)$, with $\|\xi - (s(t) + \varepsilon n(s(t))\| < \|(s(t) - (s(t) + \varepsilon n(s(t))\| = \varepsilon$. But this contradicts the fact that $s(t)$ is the point on Θ closest to $s(t) + \varepsilon n(s(t))$.

If $t_j < t$ for all j, then the same argument carries through with $s(t) + \varepsilon n(s(t))$ replaced by $s(t) - \varepsilon n(s(t))$. □

Denote the tangent hyperplane on Θ in the point ξ by $\vartheta(\xi)$, i.e., $\vartheta(\xi) = \xi + \tau(\xi) = \{\xi + b : b \in \tau(\xi)\}$.

Proposition A.2. *For any $\xi \in \Theta$ we can find $r > 0$ such that for any $x \in \vartheta(\xi)$ with $\|x - \xi\| < r$ we have that the line segment $\overline{x - \varepsilon n(\xi), x + \varepsilon n(\xi)}$ has precisely one intersection with Θ.*

Proof. We can locally parametrize Θ by a function on an open environment V of ξ in the tangent hyperplane $\vartheta(\xi)$. That is, there is an open interval $I \subseteq \mathbb{R}$ and a C^2-function $\hat\psi : V \longrightarrow I$ such that every point $z \in \{\xi + b + yn(\xi) : b \in V, y \in I\}$ can be uniquely written as $z = \xi + b + \hat\psi(b)n(\xi)$. Since $\xi \in \vartheta(\xi)$ and thus $\hat\psi(\xi) = 0$, we may assume that $I = (-\zeta, \zeta)$ for some $0 < \zeta < \varepsilon$. Choose some r such that $0 < r < \sqrt{\varepsilon^2 - (\varepsilon - \zeta)^2}$ and such that for all $x \in \vartheta(\xi)$ we have $x \in V$ whenever $\|x - \xi\| < r$.

Now if $x \in \vartheta(\xi)$ with $\|x - \xi\| < r$, then precisely one point of Θ lies on the line segment $\overline{x - \zeta n(\xi), x + \zeta n(\xi)}$. But there is no point of Θ on the line segment $\overline{x + \zeta n(\xi), x + \varepsilon n(\xi)}$, since this is entirely contained in the open ball $B_\varepsilon(\xi + \varepsilon n(\xi))$, which by the unique closest point property for $\xi + \varepsilon n(\xi)$ does not contain any point of Θ.

By the same reasoning $\overline{x - \zeta n(\xi), x - \varepsilon n(\xi)} \cap \Theta = \emptyset$. □

Proposition A.3. *Let $\varepsilon_1 < \varepsilon$. Then for any $y \in \mathbb{R}^d$ there exists a point $\hat{y} \in \mathbb{R}^d$ with $d(\hat{y}, \Theta) \geq \varepsilon_1$ and $\|y - \hat{y}\| \leq \varepsilon_1$.*

Proof. If $d(y, \Theta) \geq \varepsilon_1$, then set $\hat{y} = y$.

Otherwise, there is a unique closest point $p(y) \in \Theta$. Set

$$\hat{y} = \begin{cases} p(y) + \varepsilon_1 n(p(y)) & \text{if } n(p(y)) \cdot (y - p(y)) > 0 \\ p(y) - \varepsilon_1 n(p(y)) & \text{if } n(p(y)) \cdot (y - p(y)) < 0 \,. \end{cases}$$

Then $\|y - \hat{y}\| \leq \varepsilon_1$ is obvious, and $d(\hat{y}, \Theta) \geq \varepsilon_1$ by the unique closest point property. □

We can now modify the straight line from x to y to get a continuous curve, which is not much longer than $\|y - x\|$, but has only finitely many intersections with Θ.

For what follows, let $\alpha \in (0, 1)$ and for $0 < \delta < \varepsilon$ set $\varepsilon_1 = \varepsilon - \sqrt{\varepsilon^2 - \delta^2}$.

We construct a sequence $(\gamma_k)_{k \in \mathbb{N}_0}$ of continuous curves of finite length which becomes stationary after finitely many steps, i.e., there exists k_0 such that $\gamma_k = \gamma_{k_0}$ for all $k \geq k_0$.

Furthermore, γ_{k_0} will have only finitely many intersections with Θ and it will be only slightly longer than $\|x - y\|$, see (14).

Set $\gamma_0 = s$.

Step 1: *If* $H(s \cap \Theta) = \emptyset$, then set $\gamma_1 = \gamma_0$.

Otherwise proceed as follows: According to Proposition A.3 there exists a point $\hat{y}$ with $d(\hat{y}, \Theta) \geq \varepsilon_1$ and $\|y - \hat{y}\| \leq \varepsilon_1$. Define γ_1 as the concatenation of the lines $\overline{x, \hat{y}}$ and $\overline{\hat{y}, y}$. We have $\ell(\gamma_1) \leq \|y - x\| + 2\varepsilon_1$, and there is at most one intersection of $\overline{\hat{y}, y}$, the second line segment, with Θ, due to Assumption 3.3. Set $x_1 = x$.

After step 1 we have constructed a polygonal curve γ_1 such that $\ell(\gamma_1) \leq \|y - x\| + 2\varepsilon_1$. If γ_1 has infinitely many intersections with Θ, then all but finitely many are contained in a single line segment, $s_1 = \overline{x_1, \hat{y}}$, which satisfies $\ell(s_1) = \|\hat{y} - x_1\| = \|\hat{y} - x\| = \|(y - x) + (\hat{y} - y)\| \leq \|y - x\| + \varepsilon_1$.

Now we enter an iteration procedure. Suppose that after $k \geq 1$ steps we have constructed a polygonal curve γ_k, with the properties that $\ell(\gamma_k) \leq \|y - x\| + 2k\varepsilon_1$, and such that either γ_k has finitely many intersections with Θ, or all intersections are contained in a single line segment, $s_k = \overline{x_k, \hat{y}}$, which satisfies $\ell(s_k) \leq \|y - x\| - (k-1)(\alpha\delta - \varepsilon_1) + \varepsilon_1$.

Then we construct γ_{k+1} from γ_k as follows:

Step $k + 1$: *If* $H(\gamma_k \cap \Theta) = \emptyset$, then set $\gamma_{k+1} = \gamma_k$.

Otherwise, $H(\gamma_k \cap \Theta)$ is contained in the line segment $\overline{x_k, \hat{y}}$. Parametrize this segment by $s_k(t) = x_k + t\|\hat{y} - x_k\|^{-1}(\hat{y} - x_k)$, $t \in [0, \|\hat{y} - x_k\|]$ and let $H_k = H(\{t : s_k(t) \in \Theta\})$.

Set $t_k = \min H_k$, and let $n_k = n(s_k(t_k))$. If t_k is isolated from the left, or if $t_k = 0$, then set $r_k = 0$. Now consider the case where t_k is not isolated from the left. By Proposition A.1, s_k lies in the tangent hyperplane $\vartheta(s_k(t_k)) = s_k(t_k) + \tau(s_k(t_k))$ and we can find a small ball with radius $r_k > 0$ such that, for any t with $|t - t_k| < r_k$, the line segment $\overline{s_k(t) - \varepsilon_1 n_k, s_k(t) + \varepsilon_1 n_k}$ has at most one intersection with Θ, by Proposition A.2.

Consider the line segment $\overline{s_k(t_k - r_k) + \varepsilon_1 n_k, s_k(t_k + \alpha\delta) + \varepsilon_1 n_k}$.

If the intersection of this with the plane through $\hat{y}$, which is orthogonal to the line segment, is non-empty, denote the unique intersection point by z_k.

Then we construct γ_{k+1} as the concatenation of the following line segments:

- $\overline{s_k(0), s_k(t_k - r_k)}$, which by definition of t_k and r_k has only finitely many intersections with Θ;
- $\overline{s_k(t_k - r_k), s_k(t_k - r_k) + \varepsilon_1 n_k}$, which has at most one intersection with Θ by the construction of r_k and Proposition A.2;
- $\overline{s_k(t_k - r_k) + \varepsilon_1 n_k, z_k}$, which is completely contained in $B_\varepsilon(s_k(t_k) + \varepsilon n_k)$, which does not contain any point of Θ by the unique closest point property for $s_k(t_k) + \varepsilon n_k$;
- $\overline{z_k, \hat{y}}$, which has no intersection with Θ, because as $\|z_k - \hat{y}\| = \varepsilon_1$, there is no intersection strictly between z_k and $\hat{y}$, and z_k lies in the closure of $B_{\varepsilon_1}(\hat{y})$ (this is where we need Step 1);
- $\overline{\hat{y}, y}$.

In this case the curve γ_{k+1} has only finitely many intersections with Θ and $\ell(\gamma_{k+1}) = \ell(\gamma_k) + 2\varepsilon_1 \leq \|y - x\| + 2(k+1)\varepsilon_1$.

Otherwise, set $x_{k+1} = s_k(t_k + \alpha\delta) + \varepsilon_1 n_k$, and construct γ_{k+1} as the concatenation of the following line segments:

- $\overline{s_k(0), s_k(t_k - r_k)}$, which by definition of t_k and r_k has only finitely intersections with Θ;
- $\overline{s_k(t_k - r_k), s_k(t_k - r_k) + \varepsilon_1 n_k}$, which has at most one intersection with Θ by the construction of r_k and Proposition A.2;
- $\overline{s_k(t_k - r_k) + \varepsilon_1 n_k, x_{k+1}}$, which is completely contained in $B_\varepsilon(s_k(t_k) + \varepsilon n_k)$, which does not contain any point of Θ by the unique closest point property for $s_k(t_k) + \varepsilon n_k$;
- $s_{k+1} := \overline{x_{k+1}, \hat{y}}$, which still may have infinitely many intersections with Θ;
- $\overline{\hat{y}, y}$.

Again we have that $\ell(\gamma_{k+1}) \leq \ell(\gamma_k) + 2\varepsilon_1 \leq \|y - x\| + 2(k+1)\varepsilon_1$. Note that

$$\begin{aligned} \ell(s_{k+1})^2 &= \|x_{k+1} - \hat{y}\|^2 \\ &= \|s_k(t_k + \alpha\delta) + \varepsilon_1 n_k - \hat{y}\|^2 \\ &= \|s_k(t_k + \alpha\delta) - \hat{y}\|^2 + \varepsilon_1^2 \\ &= (\|s_k(t_k) - \hat{y}\| - \alpha\delta)^2 + \varepsilon_1^2 \,. \end{aligned}$$

In particular, $\|x_{k+1} - \hat{y}\| \leq \big|\|s_k(t_k) - \hat{y}\| - \alpha\delta\big| + \varepsilon_1 = \|s_k(t_k) - \hat{y}\| - \alpha\delta + \varepsilon_1$. Note that $\|s_k(t_k) - \hat{y}\| - \alpha\delta \geq 0$, since otherwise the line segment $\overline{s_k(t_k - r_k) + \varepsilon_1 n_k, s_k(t_k + \alpha\delta) + \varepsilon_1 n_k}$ would intersect the hyperplane orthogonal to s_k and passing through $\hat{y}$.

Thus $\|x_{k+1} - \hat{y}\| \leq \|x_k - \hat{y}\| - \alpha\delta + \varepsilon_1 \leq \|x - y\| - k(\alpha\delta - \varepsilon_1) + \varepsilon_1$.

Thus, after step $k+1$ we have constructed a polygonal curve γ_{k+1} such that $\ell(\gamma_{k+1}) \leq \|y - x\| + 2(k+1)\varepsilon_1$. If γ_{k+1} has infinitely many intersections with Θ, then all but finitely many are contained in a single line segment, $s_{k+1} = \overline{x_{k+1}, \hat{y}}$, and $\ell(s_{k+1}) \leq \|x - y\| - k(\alpha\delta - \varepsilon_1) + \varepsilon_1$.

So finally we have constructed a sequence $(\gamma_k)_{k \in \mathbb{N}_0}$ with

- $\ell(\gamma_k) \leq \|x - y\| + 2k\varepsilon_1$;
- γ_k either has only finitely many intersections with Θ, or all but finitely many intersections are contained in a segment of length at most $\|x - y\| - (k-1)(\alpha\delta - \varepsilon_1) + \varepsilon_1$.

Since $\delta < \varepsilon$, we have that $\varepsilon_1 = \varepsilon - \sqrt{\varepsilon^2 - \delta^2} = \varepsilon\left(1 - \sqrt{1 - \left(\frac{\delta}{\varepsilon}\right)^2}\right) < \frac{\delta^2}{\varepsilon}$, such that

$$\alpha\delta - \varepsilon_1 > \delta\left(\alpha - \frac{\delta}{\varepsilon}\right) > 0\,.$$

With this, and since $\|x-y\| - (k-1)(\alpha\delta - \varepsilon_1) + \varepsilon_1 \geq \ell(s_k) \geq 0$, the iteration can have at most

$$k \leq 1 + \frac{\|x-y\| + \varepsilon_1}{2(\alpha\delta - \varepsilon_1)} < 1 + \frac{\|x-y\| + \varepsilon_1}{2\delta\left(\alpha - \frac{\delta}{\varepsilon}\right)} < 1 + \frac{\|x-y\| + \varepsilon}{2\delta\left(\alpha - \frac{\delta}{\varepsilon}\right)}$$

steps before the sequence becomes stationary, and thus there exists a k_0 such that γ_{k_0} has at most finitely many intersections with Θ.

For the length of γ_k for $k \geq k_0$ we have

$$\ell(\gamma_k) \leq \|x-y\| + 2k\varepsilon_1 \leq \|x-y\| + \left(2\delta + \frac{\|x-y\| + \varepsilon}{\alpha - \frac{\delta}{\varepsilon}}\right)\frac{\delta}{\varepsilon}\,. \tag{14}$$

This can be made as close to $\|x-y\|$ as we desire by making δ small. Thus the proof is finished.

Acknowledgements

The authors thank Bert Jüttler for valuable discussions regarding questions related to differential geometry.

G. Leobacher is supported by the Austrian Science Fund (FWF): Project F5508-N26, which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

M. Szölgyenyi is supported by the Vienna Science and Technology Fund (WWTF): Project MA14-031. A part of this paper was written while M. Szölgyenyi was member of the Department of Financial Mathematics and Applied Number Theory, Johannes Kepler University Linz, 4040 Linz, Austria. During this time, M. Szölgyenyi was supported by the Austrian Science Fund (FWF): Project F5508-N26, which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

References

[1] A. Berkaoui. Euler Scheme for Solutions of Stochastic Differential Equations with Non-Lipschitz Coefficients. *Portugaliae Mathematica*, 61(4):461–478, 2004.

[2] P. Étoré and M. Martinez. Exact Simulation for Solutions of One-Dimensional Stochastic Differential Equations Involving a Local Time at Zero of the Unknown Process. *Monte Carlo Methods and Applications*, 19(1):41–71, 2013.

[3] P. Étoré and M. Martinez. Exact Simulation for Solutions of One-Dimensional Stochastic Differential Equations with Discontinuous Drift. *ESAIM: Probability and Statistics*, 18:686–702, 2014.

[4] R. L. Foote. regularity of the Distance Function. *Proceedings of the American Mathematical Society*, 92(1):153–155, 1984.

[5] I. Gyöngy. A Note on Euler's Approximation. *Potential Analysis*, 8:205–216, 1998.

[6] N. Halidias and P. E. Kloeden. A Note on the Euler–Maruyama Scheme for Stochastic Differential Equations with a Discontinuous Monotone Drift Coefficient. *BIT Numerical Mathematics*, 48(1): 51–59, 2008.

[7] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong Convergence of an Explicit Numerical Method for SDEs with Nonglobally Lipschitz Continuous Coefficients. *The Annals of Applied Probability*, 22(4):1611–1641, 2012.

[8] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1991.

[9] P. E. Kloeden and E. Platen. *Numerical Solutions of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Verlag, Berlin-Heidelberg, 1992.

[10] A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak Approximation Errors for Stochastic Differential Equations with Non-Regular Drift. 2013. Preprint, Inria, hal-00840211.

[11] S. G. Krantz and H. R. Parks. Distance to C^k Hypersurfaces. *Journal of Differential Equations*, 40:116–120, 1981.

[12] G. Leobacher and M. Szölgyenyi. A Numerical Method for SDEs with Discontinuous Drift. *BIT Numerical Mathematics*, 2015. To appear.

[13] G. Leobacher, M. Szölgyenyi, and S. Thonhauser. Bayesian Dividend Optimization and Finite Time Ruin Probabilities. *Stochastic Models*, 30(2):216–249, 2014.

[14] G. Leobacher, M. Szölgyenyi, and S. Thonhauser. On the Existence of Solutions of a Class of SDEs with Discontinuous Drift and Singular Diffusion. *Electronic Communications in Probability*, 20(6):1–14, 2015.

[15] M. Ruzhansky and M. Sugimoto. On Global Inversion of Homogeneous Maps. *Bulletin of Mathematical Sciences*, 5(1):13–18, 2015.

[16] S. Sabanis. A Note on Tamed Euler Approximations. *Electronic Communications in Probability*, 18(47):1–10, 2013.

[17] A. A. Shardin and M. Szölgyenyi. Optimal Control of an Energy Storage Facility Under a Changing Economic Environment and Partial Information. 2015. Submitted.

[18] A. A. Shardin and R. Wunderlich. Partially Observable Stochastic Optimal Control Problems for an Energy Storage. 2015. Submitted.

[19] M. Szölgyenyi. Dividend Maximization in a Hidden Markov Switching Model. 2015. Submitted.

[20] A. YU. Veretennikov. On Strong Solutions and Explicit Formulas for Solutions of Stochastic Integral Equations. *Mathematics of the USSR Sbornik*, 39(3):387–403, 1981.

[21] A. YU. Veretennikov. On the Criteria for Existence of a Strong Solution of a Stochastic Equation. *Theory of Probability and its Applications*, 27(3), 1982.

[22] A. YU. Veretennikov. On Stochastic Equations with Degenerate Diffusion with Respect to Some of the Variables. *Mathematics of the USSR Izvestiya*, 22(1):173–180, 1984.

[23] A. K. Zvonkin. A Transformation of the Phase Space of a Diffusion Process that Removes the Drift. *Mathematics of the USSR Sbornik*, 22(129):129–149, 1974.

G. Leobacher
Department of Financial Mathematics and Applied Number Theory, Johannes Kepler University Linz,

Altenbergerstraße 69, 4040 Linz, Austria
gunther.leobacher@jku.at

M. Szölgyenyi ✉
Institute of Statistics and Mathematics, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria
michaela.szoelgyenyi@wu.ac.at